HEC MONTRÉAL
École affiliée à l'Université de Montréal

Three Essays on the Role and Impacts of Data-Enabled Process Virtualization

par
Jinglu Jiang

Thèse présentée en vue de l'obtention du grade de Ph. D. en administration
(option Technologies de l'information)

Juillet 2019

HEC MONTRÉAL
École affiliée à l'Université de Montréal

Cette thèse intitulée :

Three Essays on the Role and Impacts of Data-Enabled Process Virtualization:

Présentée par :

Jinglu Jiang

a été évaluée par un jury composé des personnes suivantes :

Camille Grange
HEC Montréal
Présidente-rapporteuse

Ann-Frances Cameron
HEC Montréal
Directrice de recherche

Suzanne Rivard
HEC Montréal
Membre du jury

Kathryn Brohman
Queen's University
Examinatrice externe

Guy Paré
HEC Montréal
Représentant du directeur de HEC Montréal

Résumé

Cette thèse de trois essais est motivée par les changements et les tendances découlant de la multitude de transformations numériques récentes. Avec la mise en œuvre des technologies numériques, nous virtualisons de plus en plus les processus (c.-à-d. en supprimant les interactions physiques entre les personnes ou entre les personnes et les objets) et numérisons l'information (c.-à-d. en convertissant les représentations de notre corps, des événements, activités, interactions et environnement en bits et octets). Cette thèse cherche à comprendre les processus virtualisés dans une variété de contextes, y compris comment divers éléments du processus virtualisé peuvent modifier les interactions humaines (et parfois les actions des machines) et leurs impacts associés.

Le premier essai porte sur un processus de santé virtualisé et examine l'autosurveillance des maladies chroniques basée sur les technologies d'information (AMCTI) et ses répercussions connexes sur l'utilisation de la technologie et les résultats pour la santé. Une revue de la littérature systématique à grande échelle est effectuée afin d'élaborer un cadre théorique holistique qui capture les multiples facettes de la nature d'AMCTI. Le deuxième essai porte sur les processus virtualisés sur des plateformes numériques multifacettes. En adoptant une approche basée sur des mécanismes, cette étude théorise la façon dont les représentations des offres de plateformes numériques affectent les actions de l'agent (y compris les actions des agents humains et des agents machines sur les plateformes numériques), ce qui, à son tour, donne lieu à des résultats collectifs. Le méta-schéma proposé fournit un canevas pour les recherches futures visant à élaborer des théories contextuelles sur les interactions virtualisées entre les humains et les machines sur les plateformes numériques. Le troisième essai porte sur la prestation des services médicaux professionnels virtualisés par l'exploration des données massives sur le comportement des consommateurs dans le contexte de la consultation médicale en ligne. Une approche d'apprentissage machine est appliquée afin de développer un système des fonctionnalités liées aux services qui peuvent aider les plateformes médicales et les médecins sur la plateforme à identifier les services à forte valeur ajoutée et à promouvoir le paiement de services premium.

Mots clés: Autosurveillance informatisée, autogestion des maladies chroniques, revue de la littérature, théorie de l'actualisation des affordances, plateforme numérique, algorithme, agents machines, mécanisme, construction de théorie, consultation médicale en ligne, service de santé numérique, apprentissage machine, classification, qualité des services de santé, réputation du médecin.

Méthodes de recherche : revue de la littérature, élaboration de théories, apprentissage machine.

Abstract

This three-essay thesis is motivated by the shifts and trends arising from the multitude of recent digital transformations. With the implementation of digital technologies, we are increasingly virtualizing processes (i.e., removing the physical interactions between people or between people and objects) and digitizing information (i.e., converting representations of our body, events, activities, interactions and environment into bits and bytes). This thesis seeks to understand virtualized processes in a variety of contexts, including how various elements of the virtualized process may change human interactions (and possibly machine actions) and their associated impacts.

The first essay focuses on a virtualized healthcare process, examining IT-enabled self-monitoring for chronic disease self-management (ITSM) and its associated impacts on technology use and health-related outcomes. A large-scale systematic literature review is conducted to develop a holistic theoretical framework which captures the multifaceted nature of ITSM. The second essay focuses on virtualized processes on multi-sided digital platforms. By adopting a mechanism-based approach, this study theorizes how the representations of digital platform offerings affect agent actions (including both human agent and machine agent actions on the digital platform), which in turn give rise to collective outcomes. The proposed meta-schema provides a canvas for future research to develop context-specific theories regarding virtualized interactions among humans and machines on digital platforms. The third essay focuses on virtualized professional medical service delivery by mining massive consumer behavior data in the context of online medical consultation. A machine learning approach is applied to develop a system of service-related features that can help medical platforms and physicians on the platform identify high-value services and promote premium service payment.

Keywords: IT-based self-monitoring, chronic disease self-management, literature review, affordance actualization theory, digital platform, algorithm, machine agents, mechanism, theory building, online medical consultation, digital healthcare service, machine learning, classification, healthcare service quality, physician reputation

Research methods: literature review, theory building, machine learning

Table of Contents

List of Tables and Figures

Tables

Figures

List of Acronyms

SM	Self-monitoring
IT	Information technology
IS	Information system(s)
ITSM	IT-enabled self-monitoring for chronic care
PA	Physical activities
AI	Artificial intelligence
eWOM	Electronic word-of-mouth
ML	Machine learning
AUC	Area under the ROC curve
ROC	Receiver operating characteristic curve
LR	Logistic regression
DT	Decision tree
GB	Gradient boost tree
RF	Random forest
ADA	AdaBoost tree
XG	XGBoost tree
NB	Naïve Bayes
NN	Neutral network

Dedicated to my parents, for always supporting me.
And to my friends and mentors, for inspiring me
and guiding me on the right path.

Acknowledgments

This thesis is indebted to the support of many people. First and foremost, I would like to express my sincere gratitude to my supervisor Dr. Ann-Frances Cameron for her abiding guidance and mentoring. Many thanks, Ann-Frances! Thank you for always challenging me to step out of my comfort zone and pushing me to grow. Thank you for the patience, openness and motivation that helps me move ahead. Thank you for spending numerous hours with me every week. Thank you all for just being here for me! I could not have imagined having a better mentor for my Ph.D. journey and academic career.

My sincere thanks also go to the members of my thesis committee, Dr. Suzanne Rivard, and Dr. Anne Beaudry, for their constructive comments on the earlier versions. Suzanne, your commitment to scholarship and teaching has been such an inspiration to me! This thesis would not have been possible without excellent training from the joint program. I wish to thank Dr. Guy Paré, Dr. Liette Lapointe, Dr. Alain Pinsonneault, and Dr. Henri Barki, who let me enjoy the beauty of education and wealth of knowledge in their fantastic classes. Thank you for the immense and thought-provoking guidance that provides me with a solid foundation for my academic career! Moreover, I appreciate the assistance of Madame Line Perrier whose door has always been open when I needed help.

My time at HEC Montreal was made enjoyable in large part due to my fellows and friends, Don Maclean, Tanya Giannelia, Reza Ghaffari, Changan Zhan, Yihan Wang and Junyi Yang. Thank you for the stimulating discussions, unconditional support, and all the fun we have had for the past five years.

Lastly, I would like to thank my family for all selfless love and encouragement. Everything I am today is all for the sacrifices you made for me. Thank you for raising me and supporting me in all my pursuits. Thank you for always taking interest and pride in my work. I have a heart full of love for both of you!

Preface

A version of Chapter 1- Essay 1 has been accepted at MIS Quarterly, co-authored with Dr. Ann-Frances Cameron. I was responsible for all major areas of conceptual development, literature screening, coding and analysis, as well as the composition of the initial draft. Dr. Ann-Frances Cameron was the supervisor on this project and was involved throughout the development and manuscript writing process.

Essay 2 has been prepared for submission to a journal, co-authored with Dr. Suzanne Rivard and Dr. Ann-Frances Cameron. I was responsible for literature review, theory development and manuscript composition. Dr. Ann-Frances Cameron was the supervisor on this project and was involved throughout the conceptual development and manuscript edits. Dr. Suzanne Rivard provided tremendous help and guidance on theory building and was involved throughout the manuscript development and writing.

Chapter 3- Essay 3 is unpublished. Part of the background and literature review sections are included in conference submissions (ICIS 2019 and HICSS 2020, under review). The data was collected with the assistance of Dr. Ming Yang (Associate Professor, Central University of Finance and Economics, China). I was responsible for data collection methodology design, data analysis, literature review, conceptual development and manuscript composition. Dr. Ann-Frances Cameron was the supervisor on this project and was involved throughout conceptual development and manuscript edits.

Introduction

The recent shifts and trends arising from a wide array of digital transformations are impacting every corner of our lives. New technologies, such as digital platforms, artificial intelligence, cloud techniques and virtual networks, have transformed physical activities and interactions into ones that occur in a virtual space. Through the removal of physical interactions between people or between people and objects, and digitization of information by converting representations of our body, events, activities, interactions and environment into bits and bytes, processes are becoming increasingly virtualized. As more analog information is encoded into digital forms and various physical processes are now being conducted virtually, many new questions arise. For example, how does IT enable such process virtualization? How do virtualized processes influence human (and possibly machine) agents' behaviors? What are the impacts on individuals, organizations and society?

This thesis seeks to understand the role and impacts of process virtualization, especially for those processes enabled by an exponentially growing amount of user data that is diverse and detailed. The extent to which a process can be successfully virtualized depends on the nature of a process – for example, its sensory requirements, relationship requirements, synchronism requirements and control requirements – and the IT capabilities that can proliferate the virtualization process, such as the representation capability, reach across time and space, as well as monitoring capability (Overby 2008). With the support of new technologies that are ever-connected and smart, the digital transformation we face today is far more than simply automating manual processes and encoding information into a format that can be recognized by computers. These virtualized processes often employ technologies to deepen collaboration among actors – including both human and machine actors – and these collaborations can form the basis of new ecosystems that generate profound impacts on individuals, organizations and society.

The three essays in this thesis examine context-specific virtualized processes and their impacts. Table 0.1 briefly presents the three studies, their contexts, and how they contribute to understanding virtualized processes.

Table 0.1 Summary of Three Essays

	IT	Virtualized Process	Type of Study	Results related to virtualized processes
Essay 1	Self-monitoring technologies	Chronic disease self-management	Literature review	Four ITSM affordances emerged as enablers; Four types of immediate outcomes and two ultimate health outcomes of virtualized processes emerged.
Essay 2	Multi-sided digital platforms (general purpose)	The interaction among agents and platform offerings	Theory building (conceptual)	A meta-schema with seven mechanisms which can be leveraged by future research to examine the digital ecosystem empowered by human- and machine-agent generated data and their joint actions through virtual interactions.
Essay 3	Multi-sided digital platforms (for online medical consultation)	Medical consultation service	Empirical (machine learning)	A system of service-related features extracted from massive consultation data, which can be used to help the medical platform identify high-value services that may lead to payment.

Summary of Essay 1
IT-Enabled Self-Monitoring for Chronic Disease Self-Management: An Interdisciplinary Review

Essay 1 examines the virtualized healthcare process. This study aims to understand IT-enabled self-monitoring (ITSM) for chronic disease self-management. Self-monitoring, defined as "*awareness of symptoms or bodily sensations that is enhanced through periodic measurements, recordings and observations to provide information for improved self-management*" (Wild and Garvin 2007, p. 343), is an essential component in chronic care, which has been implemented in healthcare practices for quite some time (McBain et al. 2015). With the recent development of technologies and digital platforms, typical self-monitoring activities such as self-recording of symptoms, analyzing self-recorded data and adjusting behaviors are increasingly supported by IT. Although ITSM is ongoing and practically important, the accumulation of knowledge in this area is fragmented. For example, medical research almost exclusively focuses on the implementation and effectiveness of clinical treatment with ITSM treated as a blackbox, while computer science studies focus on developing new ITSM tools with limited understanding of how ITSM is used and experienced in practice.

To address this opportunity, the primary objective of this study is to understand what has been done in the relevant fields of research and develop an overarching theoretical framework to organize the extant studies. We systematically review 159 studies published in 108 journals and conferences between 2006 and 2017. By adapting Affordance Actualization Theory, we organize the existing literature on ITSM for chronic disease management into four themes: key ITSM functionalities that enable affordances; effects on ITSM system use; effects on the achievement of chronic care goals; and the role of intermediary outcomes.

ITSM, as a virtualized chronic care process, is supported by IT with four key types of affordances: (1) *preparation* that trains and motivates patients to engage in the subsequent SM, which is often realized through IT functionalities such as education delivery and goal setting; (2) *data collection* that supports fully automated or semi-manual data entry and recording; (3) *user reflection* that helps patients understand the records and take actions, which is often supported by IT functionalities such as data display, push message and gamification; and (4) *social connection* that allows peer interactions and collaboration between patients and physicians.

The dominant stream of research is how the implementation of ITSM (usually as part of a complex healthcare intervention) impacts behavior change and health improvement. Positive associations have been found regarding physical activity, diet and weight management outcomes. However, the results are less consistent for diabetes management and improvement of self-rated quality of life. Several studies examined ITSM use frequency as one of the direct outcomes. However, whether or not ITSM use mediates the impact of ITSM on healthcare outcomes is less known. In addition, since ITSM is usually implemented as part of a complex healthcare intervention, whether the mixed results are caused by IT or another competing non-IT component is less known.

A small number of studies examined potential intermediate outcomes other than ITSM use, which may serve as important mechanisms to explain how ITSM can be effective in different circumstances. Four types of intermediate outcomes emerge: (1) patient learning and self-reflection, (2) patient-provider co-management of chronic conditions, (3) social interaction with families and peers, and (4) intervention satisfaction and compliance. These outcomes are largely supported by specific IT functionalities and self-management procedures. Thus, how to design ITSM and interventions to minimize procedural barriers can be an important research direction to facilitate the effective implementation of ITSM for chronic care.

For each abovementioned theme, we identify what is known (key consistent and inconsistent results), what is unknown, and opportunities for future research. We also discuss cross-theme opportunities for future research where more diverse theoretical perspectives can contribute to our understanding of the phenomenon.

Summary of Essay 2
Human Agents, Machine Agents and User Commitment on Digital Platforms: A Mechanism-Based Meta-schema

Essay 2 examines the virtualized social process among human agents and machine agents on digital platforms. Social interactions and transactions happening on digital platforms have attracted intensive research interest for the past decade (e.g., studies on online review, user-generated content, digital collaboration, and e-commerce). Notwithstanding the contributions of extant research in identifying the antecedents of user decisions and actions on digital platforms, existing research has been limited to humans as actors and has largely ignored the role of machine agents, despite their increasing presence on platforms (e.g., chat bots, virtual assistants, automatic transaction agents). In addition, the majority of the research adopts a variance-based theorizing approach, which does not reveal the cogs and wheels operating behind the related constructs. To address these two gaps, we adopt a mechanism-based approach to develop a meta-schema of seven mechanisms that explain how the representation of digital platform offerings affects agent actions (for both human agents and machine agents) and how those agent actions give rise to collective outcomes.

A mechanism "consists of entities and activities organized in such a way that they are responsible for the phenomenon" (Illari and Williamson 2012, p.120). The entities are the producers of changes, which can be identified by their properties that allow their engagement in the activities (Glennan 1996). A mechanism is organized such that its entities and activities are set up to do something (Craver 2001). Examining the mechanism that produces or underlies a phenomenon is to explain *why* a phenomenon happens by describing *how* some mechanisms produce the phenomenon (Halina 2017). For the context of this study, we focus on two broad types of mechanisms: social mechanisms (to explain human actions) and computational mechanisms (to explain machine actions).

Social mechanisms are defined as "*social processes having designated consequences for designated parts of the social structure*" (Merton 1968, cited in Hedström and Swedberg 1998 p.6). They explain how social-level causes generate social-level outcomes through social processes. Hedström and Swedberg (1998) posit that to explain a macro-level social phenomenon, three types of mechanisms should come into play: (1) situational mechanisms in which social context influences individual actors' beliefs, habits and cognitive frames; (2) action-formation mechanisms in which individual actors' opportunities and cognitive frames lead to or change the actors' behaviors; and (3) transformational mechanisms in which individual actors' actions produce social patterns and outcomes.

Computational mechanisms, on the other hand, are responsible for the functions and behaviors of computing systems by describing how "the input and output information streams are causally linked … along with the specific structure of information processing" (Miłkowski 2014, p.221). Machine agents on digital platforms operate based on pre-designed algorithms. Since the inputs are largely influenced by human actions, and the outputs influence human agents' subsequent decision making and actions, it is important to understand how human agents and machine agents jointly produce commitment and non-commitment actions, which may further give rise to collective outcomes that go beyond a single agent.

By analyzing the key attributes and action trajectory of human and machine agents in the context of multi-sided digital platform, we propose seven mechanisms and the associated propositions that explain the relationships between digital platform offerings, agent actions, commitment and non-committal outcomes at agent level, and the collective outcomes:

- ***Proposition 1a***: *Human agents form cognitive frames (e.g., beliefs, attitudes, perceptions, motivations, desires, preferences, affects) through observational learning or direct interaction with the offering representation, which is a sense-making process to understand the offering's causal capacities.* [**Mechanism 1a:** human agent cognitive frame]

- ***Proposition 1b***: *Machine agents' ability to build up knowledge of the environment (including knowledge of the offering and its surrounding social dynamics) varies according to their percepts collection and processing strategies.* [**Mechanism 1b**: machine agent belief formation]

- ***Proposition 2a***: *Human agents' cognitive frame compositions influence action selection, and different patterns of the cognitive formation may have different impacts on actions.* [**Mechanism 2a**: human agent action formation]

- ***Proposition 2b***: *Machine agents' action formation is influenced by their decision-making architecture (e.g., logic-deduction, reactive, belief-desire-intention, layered), so that even with the same percept and goals, the action plan selection may be different.* [**Mechanism 2b**: machine agent action selection]

- ***Proposition 3***: *Offering causal capacity update is a joint process through which human and machine agent actions provide new information, and the platform's general-purpose algorithms selectively use this information to update the offering's attributes.* [**Mechanism 3**: causal capacity update]

- ***Proposition 4***: *The platform's general-purpose algorithm selects information from the actual offering causal capacity to update the offering representation,* *depending on its decision-making architecture.* [**Mechanism 4**: offering representation update]

- ***Proposition 5**: The nature of individual commitment and non-committal actions will influence the emergence of collective outcomes, which will vary from composition (e.g., linear addition of individual transactions) to compilation (e.g., maximum bidding results, networked negotiation).* [**Mechanism 5**: collective outcome emergence]

By simultaneously considering social mechanisms and computational mechanisms, the meta-schema with seven mechanisms provides a vocabulary and constitutes a canvas that can be leveraged by future research. At the end of the essay, evaluation criteria and potential approaches that can be used for further context-specific theory building are discussed.

Summary of Essay 3
Which Physicians Attract Paying Customers? Mining Massive Service Data to Understanding Patient Payment in Freemium-Based Online Medical Consultation

Essay 3 examines virtualized medical service delivery on multi-sided digital platforms. Over the last decade, a new wave of healthcare service digitization has emerged including new digital products (e.g., mobile apps for healthcare), healthcare channels (e.g., online healthcare communities, digitally delivered healthcare diagnoses and treatment recommendations), and business models (e.g., the involvement of digital health companies). This study focuses on one particular type of digital healthcare service – online medical consultation on multi-sided digital platforms.

Previous studies on online medical consultation have investigated various individual and contextual factors (e.g., physician reputation, patients' feedback and patient-physician interaction) that are associated with patient selection of physicians and payment, which contribute to a physician's success on the platform. Whereas these studies provide useful managerial implications, they examine a small set of factors in isolation. In addition, these studies do not take into account the specifics of the business model that define the value propositions and revenue generation mechanisms of the digital platform. Additional research is needed to develop a more holistic understanding of the factors that contribute to patients' selection of physicians and payment, as well as an understanding that takes into account the characteristics of the digital platform's business model. As a first step, this study focuses on online medical digital platforms employing freemium (FREE+ preMIUM), a business model with a multi-tiered pricing strategy that allows the coexistence of free (i.e., the free trials) and paid versions (i.e., the premium) of goods and services (Kumar 2014; Wagner et al. 2014). We aim to understand (1) the key service-related features that are associated with patient payment, as opposed to free-trial only appointments, in freemium-based online medical consultation, (2) the relative importance of these features, and (3) how these features interact, linearly or non-linearly, in relation with premium payment.

A machine learning approach is used to investigate the research questions by mining 1.5 million consultation records from a large multi-sided medical consultation platform in China. Previous studies on freemium (and the relevant research on sampling and versioning) have revealed four types of intermediate mechanisms that may contribute to premium payment – consumer awareness and learning, consumer valuation, product/service quality and consumer involvement. Inspired by these mechanisms, as well as the antecedents that have been examined in online medical consultation literature, we extract 18 service-related features that are potentially relevant to informing premium payment. After performing the feature selection procedures, 11 features are retained to form a model for further machine learning tasks – solving a binary classification problem by learning the functions that can map the system of inputs (i.e., the 11 features) to the output (paid vs. free services). Eight machine learning algorithms are used to cross-validate model performance and the importance of the selected features. The results support the usefulness of the model. However, as opposed to previous studies that highlighted the key role of physician reputation, the results show that although physician reputation is important, service quality and patient involvement appear to contribute more to the premium payment.

This study contributes to online medical consultation by developing a holistic model with a system of service-related features to help the platform target high-value service providers and the type of services that attract premium payment. Methodologically, our machine learning approach complements previous regression-based analysis by effectively mining a massive amount of fine-grained consumer behavior data with a high number of dimensions, and providing insights on complex interactions among the relevant service features that contribute to payment. Platform owners and administrators can utilize our model to identify and promote high-value services and service providers, which in turn should help support the long-term success of the platform.

References

Craver, C. F. 2001. "Role Functions, Mechanisms, and Hierarchy," *Philosophy of science* (68:1), pp. 53-74.

Glennan, S. S. 1996. "Mechanisms and the Nature of Causation," *Erkenntnis* (44:1), pp. 49-71.

Halina, M. 2017. "Mechanistic Explanation and Its Limits," *The Routledge Handbook of Mechanisms and Mechanical Philosophy.* Routledge, London, pp. 213-224.

Hedström, P., and Swedberg, R. 1998. *Social Mechanisms: An Analytical Approach to Social Theory.* Cambridge University Press.

Kumar, V. 2014. "Making" Freemium" Work," *Harvard business review* (92:5), pp. 27-29.

Illari, P. M., and Williamson, J. 2012. "What Is a Mechanism? Thinking About Mechanisms across the Sciences," *European Journal for Philosophy of Science* (2:1), pp. 119-135.

McBain, H., Shipley, M., and Newman, S. 2015. "The Impact of Self-Monitoring in Chronic Illness on Healthcare Utilisation: A Systematic Review of Reviews," *BMC Health Services Research* (15:565), p. 1-10 (doi: 10.1186/s12913-015-1221-5).

Merton, R. K. 1968. *Social Theory and Social Structure.* New York: Free Press.

Miłkowski, M. 2014. "Computational Mechanisms and Models of Computation," *Philosophia Scientiæ. Travaux d'histoire et de philosophie des sciences* (18:3), pp. 215-228.

Overby, E. 2008. "Process Virtualization Theory and the Impact of Information Technology," *Organization science* (19:2), pp. 277-291.

Wagner, T. M., Benlian, A., and Hess, T. 2014. "Converting Freemium Customers from Free to Premium—the Role of the Perceived Premium Fit in the Case of Music as a Service," *Electronic Markets* (24:4), pp. 259-268.

Wilde, M. H., and Garvin, S. 2007. "A Concept Analysis of Self-Monitoring," *Journal of Advanced Nursing* (57:3), pp. 339-350.

Chapter 1- Essay 1

IT-Enabled Self-Monitoring for Chronic Disease Self-Management: An Interdisciplinary Review

Jinglu Jiang
Department of Information Technologies, HEC Montreal
3000, chemin Côte-Sainte-Catherine
Montréal, QC
H3T 2A7
CANADA
jinglu.jiang@hec.ca

Ann-Frances Cameron
Department of Information Technologies, HEC Montreal
3000, chemin Côte-Sainte-Catherine
Montréal, QC
H3T 2A7
CANADA
ann-frances.cameron@hec.ca

* This paper has been accepted at MIS Quarterly

Abstract

Self-monitoring is a strategy that patients use to manage their chronic disease and chronic disease risk factors. Technological advances such as mobile apps, web-based tracking programs, sensing devices, wearable technologies, and insideable devices enable IT-based self-monitoring (ITSM) for chronic disease management. Since ITSM is multidisciplinary in nature and our understanding is fragmented, a systematic examination of the literature is performed to build a holistic understanding of the phenomenon. We review 159 studies published in 108 journals and conferences between 2006 and 2017. By adapting Affordance Actualization Theory, we develop an overarching framework to organize the existing literature on ITSM for chronic disease management. Four themes emerge: key ITSM functionalities that enable affordances; effects on ITSM system use; effects on the achievement of chronic care goals; and the role of intermediary outcomes. For each theme, we identify what is known, what is unknown, and opportunities for future research. We also discuss cross-theme opportunities for future research where more diverse theoretical perspectives can contribute to our understanding of the phenomenon. This work provides research directions for IS researchers studying ITSM for chronic disease self-management.

Keywords: IT-based self-monitoring, chronic disease self-management, literature review, ITSM use, affordance actualization theory

1.1 Introduction

Chronic diseases, such as diabetes, obesity and asthma, are among the most prevalent and costly health problems worldwide (Bauer et al. 2014, WHO 2014). Chronic diseases are also highly preventable, and many of them share common risk factors, such as lack of exercise, nutrition deficiency, being overweight, smoking, and excessive drinking (CDC 2016). Mitigating these risk factors is key for chronic disease management. For patients with chronic disease(s), chronic disease management involves using complex combinations of strategies to manage the disease(s) so as to slow progression or to manage the high-risk factors associated with chronic health conditions (Mallery and Rockwood 1992; WHO 2007). These strategies aim to help patients manage their chronic health conditions on a day-to-day basis. They include both clinical interventions as well as home-based self-management which encourages the involvement of individuals and their families in their own care (Martin 2007; WHO n.d.).

Self-monitoring (SM) is often considered an essential component in chronic disease self-management (shortened to "chronic care" hereinafter) and patients' willingness to self-monitor largely affects the achievement of positive health outcomes (Huygens et al. 2017, McBain et al. 2015). Self-monitoring is the "*awareness of symptoms or bodily sensations that is enhanced through periodic measurements, recordings and observations to provide information for improved self-management*" (Wild and Garvin 2007, p. 343). SM involves self-recording of symptoms and behaviors, interpreting the self-recorded data, adjusting behaviors accordingly, and applying treatments or seeking professional help as a result of self-awareness (Epstein et al. 2008, McBain et al. 2015). Historically, most SM systems were paper-based and memory-based, and while useful for patient empowerment, various problems—such as low compliance, recall bias, and difficulties in tracking moment-to-moment information—hindered their potential effectiveness (Faurholt-Jepsen et al. 2016).

The recent growing interest in IT-based self-monitoring for chronic care (ITSM) potentially overcomes the difficulties of the traditional systems (Chen and Yeh 2015, Faurholt-Jepsen et al. 2016, Panagioti et al. 2014). Technological advances such as mobile apps, affordable sensing devices, wearables, insideable technologies (e.g., in-body implants, under-skin sensors, or ingestible smart pills) improve the capabilities of SM-based healthcare programs. For example, the interactive visualization of ITSM technologies helps people better understand health patterns over time (Cuttone et al. 2013). Persuasive functions such as adaptive recommendations help users connect with their health professionals at the appropriate moment (Fairburn and Rothwell 2015). Connectivity and mobility allow people to manage their information seamlessly (Grönvall and Verdezoto 2013b). The pervasiveness of ITSM is also evidenced by the increasing popularity of these tools among healthcare consumers and the general population. It is forecasted that by 2019, more than 245 million smart wearable devices will be sold, representing more than $25 billion for smartwatches and fitness trackers (CCSInsight 2017). Thus, ITSM has great potential to help people control and manage the high-risk factors related to their chronic diseases (Kennedy et al. 2012, McBain et al. 2015).

ITSM is ongoing and practically important, yet the accumulation of knowledge in this area is fragmented, and several research disciplines examining ITSM have developed (Chomutare et al. 2011, James et al. 2019; Lehto and Oinas-Kukkonen 2011, Lupton 2014). Each of the emerging research streams tends to focus on specific aspects of ITSM, and – to our knowledge – no comprehensive framework exists to tie the disparate streams of research together. For example, medical research on chronic disease mainly focuses on the implementation and effectiveness of clinical interventions with ITSM as a regular intervention component. This stream of research generally treats IT and the use of IT as a black box and largely ignores the impacts of user perceptions and experiences. Another key stream of research is found in IS and computer science, which focuses on how to design more effective and useful SM tools by understanding how SM systems are used and experienced in practice (Chung et al. 2016, Epstein 2015). This research stream examines IT in detail (Ayobi et al. 2016) but largely ignores the chronic disease context and patients' specific needs. While ITSM is multidisciplinary by nature, these different streams of research have not been woven together into a cohesive understanding of ITSM for chronic care and this failure to capture the multifaceted nature of ITSM may cause a disjointed accumulation of knowledge. Thus, a synthesis of the current research which makes connections between divergent literatures is needed in order to develop a more holistic understanding of this phenomenon and build a cumulative knowledge tradition.

To address this opportunity, the primary aim of this article is to provide a systematic cross-disciplinary synthesis of the literature that contributes to our understanding of ITSM for chronic care and provides research directions for IS researchers studying ITSM. This aim is achieved by: (1) organizing the research based on an overarching framework, specifically, affordance actualization theory (Strong et al. 2014); (2) using the framework to synthesize the results to identify what we know and do not know; and (3) identifying future research opportunities. We are not suggesting that future researchers limit themselves to the actualization affordance lens, however, our detailed framework should help researchers identify areas of interest and specify how their own research complements, replicates, or diverges from the larger body of ITSM for chronic care research.

Our synthesis of the literature makes several contributions. First, it provides an overarching theoretical framework to organize extant research. The organizing framework enables an overview of the current status of ITSM research related to chronic care from several disciplines and provides a repository of accumulated knowledge on ITSM. It also helps us surface gaps in our understanding and identifies future research directions. For example, recent technological advancements should increase ITSM use and enhance positive outcomes but are as of yet largely under-examined in the literature. User learning and social support from peers or providers are two intermediary mechanisms highlighted in practitioners' chronic care practices, which also lack research. Outcomes specific to medication and chronic conditions have received less attention given that a high proportion of studies focusing on physical activity and weight management outcomes. Second, the framework also helps us discover several key overarching issues in this field of research, namely that the research on ITSM for chronic care is largely fragmented, there is a shallow understanding of the role of IT, and there is a paucity of strong theory. Third, our overarching theoretical framework is IS-centric. It emphasizes the role of IT functionalities and highlights four sets of ITSM affordances and their associated intermediate outcomes.

1.2 Background

1.2.1 Chronic Disease Self-Management and Self-Monitoring

Chronic disease management is "an intervention designed to manage or prevent a chronic condition using a systematic approach to care and potentially employing multiple treatment modalities" (Weingarten et al. 2002, p. 2). As opposed to acute disease, chronic disease is lengthy, not curable and usually gradual, thus requiring longitudinal supervision and reciprocal knowledge between the patient and healthcare providers (Lorig 1996). Chronic disease management is a broad term that encompasses chronic disease prevention and efforts to reduce or control risk factors (Peytremann-Bridevaux and Burnand 2009). We limit our review to focus on an individual's chronic disease management, rather than chronic disease management for entire populations.

The important role of self-management in chronic care has been highlighted (Lorig et al. 1999). While managing disease was traditionally viewed as the responsibility of doctors, modern chronic disease management recognizes the importance of a strong partnership between patients, healthcare providers and families (Barr et al. 2003; Coleman et al. 2009; Wagner et al. 2001). As a result, more attention is now given to patients' self-management of chronic conditions, which requires skills such as detecting bodily symptoms, using monitoring machines, understanding measurements and communicating self-monitored information (Paterson et al. 2001).

Self-monitoring (SM) is one essential strategy for self-management (Bartholomew et al. 1993; Bodenheimer et al. 2002; Farmer et al. 2007; Norris et al. 2001). Whereas self-management is a broad term encompassing treatment adjustment, symptom management and self-motivation, SM is a more specific term that encompasses the activities necessary to track and use one's own information.[1] SM is different from providers' monitoring of the patients where the accessibility of the information on the patients' side is limited. The healthcare literature does not define SM consistently and uses multiple terms

---

[1] In cases where patients need assistance to use their own information, families, acquaintances, or trusted entities who are involved in this self-care process may also have access to this information. While important, these are outside of the scope of the current review.

interchangeably, such as self-management, self-care, symptom management, self-tracking, and self-recording (Legorreta et al. 2000; Minet et al. 2010; Schilling et al. 2002). We adopt the term SM, which includes self-recording of symptoms and behaviors, interpreting the self-recorded data, adjusting behaviors accordingly, and applying treatments or seeking professional help as a result of self-awareness (Epstein et al. 2008; McBain et al. 2015).

1.2.2 IT-enabled Self-Monitoring for Chronic Care

People may use any tool, including paper and pencil, to keep track of their information, but the recent developments in digital technologies offer new opportunities and increase the complexity of SM systems. The result has been an increase in research on ITSM in various contexts including healthcare, education, the workplace and one's personal life. We define ITSM as the use of digital technologies to enable patient SM – i.e., the use of digital technologies to support self-recording of symptoms and behaviors, interpreting the self-recorded data, and adjusting behaviors accordingly.

It is important to note that ITSM technology users, usually the patients or healthcare consumers, are both the providers and the users of the information (Marx 2002). Even when the use of ITSM is mandated by a physician as part of treatment, it is still the patients (and their personal care attendants or family members in cases where the patients need assistance) who record their own information and use this information to better manage their chronic diseases.[2] Thus, ITSM in healthcare is an emerging medical approach where the patient maintains significant control (Swan 2009). Patients usually have ultimate control over what data are entered into the ITSM, when to input the data, and whether or not they share their data with physicians, family members, or other ITSM users.

ITSM can be illustrated using Li et al.'s (2010) five-stage process of personal informatics[3]. First, patients prepare to use the system, which can include activities such as education on system use and setting goals and targets. Second, patients need to observe

[2] If it is the healthcare provider who exclusively views or uses this information, it operates more as a surveillance rather than a SM system.
[3] Li et al.'s model is for all types of ITSM and is not specific to ITSM for chronic care.

the information about themselves and record this information. Sometimes the patient needs to track the information independent from the activity that occurred (e.g. eating followed by the recording of calories), and sometimes the system automatically tracks the information (e.g. automatic tracking of steps while walking). Third, the collected data is integrated and displayed by the system for further analysis and interpretation by the patients, their acquaintances and/or healthcare professionals. Fourth, patients and healthcare professionals need to understand and reflect on the information produced by the system to discover patterns, correlations, and insights related to patients' health statuses. Lastly, patients need to act on what they have learned. Based on the information produced by the ITSM systems, patients may individually change their behaviors or work with healthcare professionals to adjust treatment of their chronic disease.

1.3 Overarching Framework: ITSM Affordance Actualization

We develop an overarching framework with which to organize the extant research by adapting Strong et al.'s (2014) affordance actualization theory. IT affordances are not simply the technology's physical characteristics, but the action possibilities permitted by the IT functionalities (Gibson 1979). The potential embedded in these affordances is realized during the actualization process, defined as "the actions taken by actors as they take advantage of one or more affordances through their use of the technology to achieve immediate concrete outcomes in support of organizational goals" (Strong et al. 2014, p.70). The actualization process includes both actions (e.g., use of IT) and the immediate outcomes of those actions (i.e., the expected immediate outcomes that are perceived as useful for achieving the ultimate goals). The link between action and the immediate outcomes is iterative with the immediate outcomes providing feedback to influence subsequent ITSM actions. These immediate outcomes are also the link between users' actions and achieving ultimate goals. Finally, there are various external factors (e.g. work environment) that support and constrain this actualization process (Strong et al. 2014).

The affordance-actualization lens is appropriate as an overarching framework for organizing the research on ITSM for chronic care for several reasons. First, this lens includes the influence of IT functionalities on both use and outcomes. Second, it has previously been used to study individual-level IT use and impacts (e.g. Anderson and

Robey 2017; Lehrig et al. 2017; Thapa and Sein 2018), as well as to study non-chronic disease SM for the general population (e.g. Mettler and Wulf 2019). Third, affordance-actualization theory highlights important links that – theoretically – should exist by focusing on *how* IT results in specific outcomes. This strong theoretical foundation on which to synthesize the diverse ITSM for chronic disease management literature helps surface gaps that are opportunities for future research.

We adapt affordance actualization theory to our ITSM context (see Figure 1.1). Strong et al. focused on *immediate concrete* outcomes, but we i) expand these to also include psychological, cognitive and affective outcomes which are relevant in the context of ITSM for chronic care, and ii) change the term to *intermediate outcomes* to reflect that some of the outcomes are not instantaneous and to put the focus on their potential role as important links between ITSM use and chronic care goal achievement. While Strong et al. took a multilevel approach and focused on achieving overarching *organizational* goals, our ITSM review focuses on an *individual's* chronic care goals. Finally, we contextualize the external factors that support and constrain the actualization process as the non-IT complementary components of a medical intervention (hereinafter abbreviated as "non-IT components") that impact the actualization process.

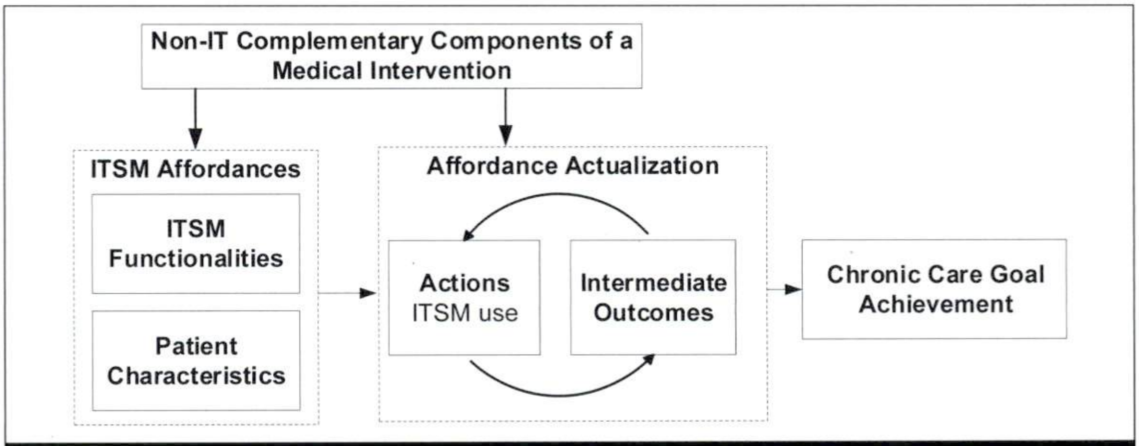


Figure 1.1 Overarching Framework for ITSM research (adapted from Strong et al. 2014)

1.4 Methodology

We mapped existing research across multiple disciplines to our overarching theoretical framework. This enabled us to synthesize extant research, develop a holistic understanding of what has been studied, and examine the theoretical foundations suggested for those relationships. It also enabled us to surface gaps and propose directions for future research.

We followed a formal systematic literature review process for searching and screening articles, which is presented in Figure 1.2 (Okoli and Schabram 2010; Webster and Watson 2002). The search strategy we adopted, while not exhaustive, included as many studies as possible. Eight digital libraries were searched: EBSCO host (including MEDLINE), ABI/INFORM, ACM digital library, ScienceDirect, IEEE Xplore, JSTOR, PsycINFO, and Web of Knowledge. Titles and abstracts (or titles and topics for Web of Knowledge) of English articles published in peer-reviewed journals and conference proceedings from 2006 to 2017 were searched using the following terms: "self-monitor*", "self-surveillance", "self-track*", "personal informatics", "personal analytics", and "electronic personal archive".

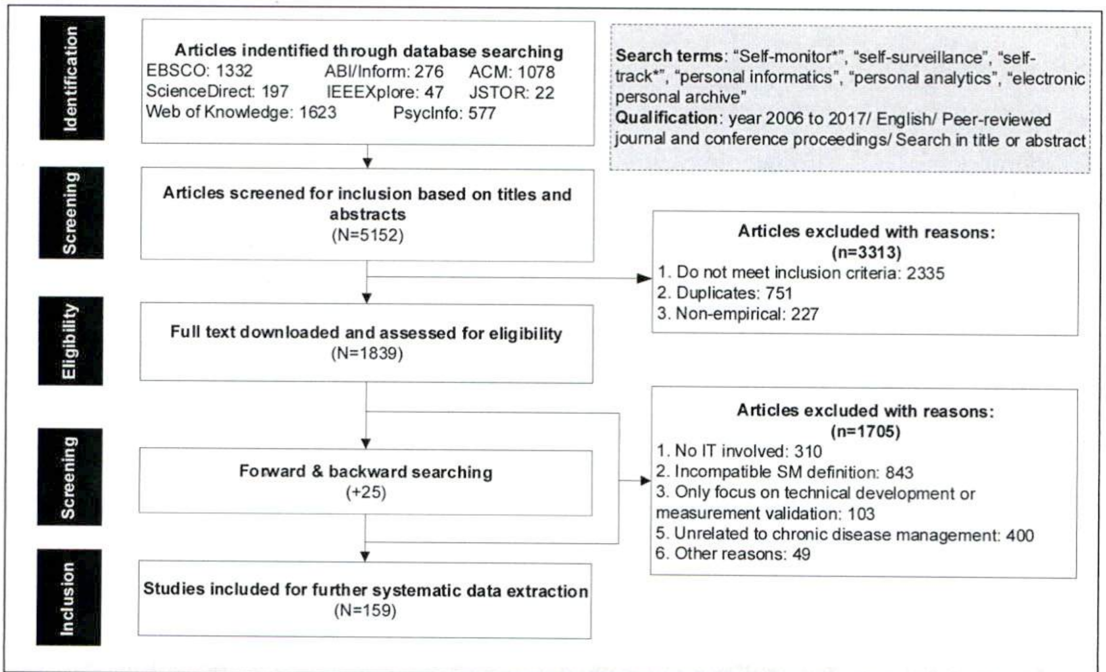


Figure 1.2 Literature Review: Searching and Screening Process

Articles were screened based on a review of titles and abstracts, with researchers reading the full text when needed. The identified studies were screened according to the inclusion and exclusion criteria presented in Table 1.1. The database search resulted in 5,152 studies. After removing the duplicates, forward and backward searching, and iteratively applying the inclusion and exclusion criteria, 159 studies remained for further analysis (articles with asterisk in the reference list).

While the affordance actualization framework was used to guide our synthesis, these studies did not explicitly examine "affordances" or use an affordance-based perspective. However, five concepts from our framework were explicitly used to analyze the literature: ITSM technological functionalities, ITSM use, intermediate outcomes, chronic care goal achievement and non-IT components of the medical intervention (i.e. external factors in the original framework). The affordance actualization theory also involves feedback loops to create an iterative process. Feedback loops were not coded as an independent concept, but the direction of relationships between constructs were captured where applicable. Each study's constructs were mapped onto these main concepts. Next, the long list of constructs in each of these concepts are distilled into a set of sub-constructs through iterative discussion between the researchers until we reached consensus. In addition, new concepts were allowed to emerge if they did not fit the affordance actualization framework.

Table 1.1 Inclusion and Exclusion Criteria for Article Screening
Inclusion criteria
1. Empirical research articles 2. Studies that examine a specific chronic disease[4], or chronic disease prevention and management for those with a chronic condition 3. Studies that examine at least one of the following aspects regarding ITSM systems: usability evaluation of new ITSM design, implementation of ITSM systems, ITSM systems adoption or post-adoption use, and the impact of using ITSM systems
Exclusion criteria

[4] The list of chronic diseases was obtained from the website for the Council for Medical Schemes (https://www.medicalschemes.com/medical_schemes_pmb/chronic_disease_list.htm).

Table 1.1 Inclusion and Exclusion Criteria for Article Screening
1. Non-empirical articles (e.g. editorials, abstracts, workshop/conference summaries, research proposals, reports based on descriptive data without examining scientific relationships and results, clinical protocols, intervention designs without testing, literature reviews, conceptual papers)
2. Not related to any chronic disease
3. No IT involved (e.g. studies which only include paper-based SM, memory-based SM or medical devices such as traditional weight scales that do not have the capacity to store, transmit, and retrieve historical information)
4. Incompatible definition of SM: patients are not allowed to use their own information (e.g. clinical self-assessment where the results are only provided to healthcare providers); self-monitoring as a personality trait that focuses on how people control their expressive behavior to accommodate social cues; firms' self-monitoring of their business performance; individuals not monitoring their "self" information (e.g. monitoring electricity consumption of a house)
5. Studies that only focus on technical development or new measurement development (e.g. hardware and algorithm improvement studies, and clinical measurement design that uses SM as a data collection method)
6. SM used only as a measurement instrument in the study
7. No primary and/or human data related to use or impacts collected (e.g. the descriptive analysis of SM app features without showing their implementation, use or impacts)
8. Studies only providing descriptive statistics without further investigating any relationships (e.g. the number of users for the SM app)

1.5 Profile of Studies and ITSM Research Trends

A general profile of the 159 studies is presented in Table 1.2, which displays publication trends by discipline, methodology, research objectives, chronic condition, and IT type. Research interest in ITSM is increasing, with over 70% of ITSM studies appearing from 2014 to 2017. The majority of studies in all three time periods were published in medical journals, with the second largest group being published in intersection journals. Almost half of the studies focused on research objectives related to medical intervention designs and evaluations (N=79), and over half employed experimental methodologies (randomized controlled trial experiments N=68, non-randomized experiment or intervention N=42), which may be unsurprising given the significant proportion of studies from the medical field. A wide range of chronic conditions are present in the studies, with the most frequent being obesity (N=53), diabetes (N=37), and psychiatric conditions (N=16). ITSM studies involving psychiatric conditions experienced a large increase in research attention, from zero studies in 2010-13 to 15 studies in 2014-17. A wide range of IT devices are also represented in the studies, with the most frequently used being mobile or tablet apps (N=58), followed by web-based SM (N=42), and medical devices

(N=36). One general trend is a move towards using smarter and more connected ITSM devices, with smart wearables becoming frequently used between 2014 and 17 (N=15) and some recent studies investigating insideables (e.g., Mathieu-Fritz et al. 2017; Polonsky et al. 2017).

While a variety of ITSM types were used for different chronic conditions (see Appendix A, Table A1), a few patterns are noticeable. Medical devices (such as glucometers) are particularly popular for diabetes while wearables (such as smart fitness trackers) are often used for obesity. ITSM for psychiatric conditions most often employs mobile or tablet apps that allow questionnaire-based SM for moods and symptoms.

Table 1.2 Profile of the Studies by Discipline and Year

	2006 - 2009				Total	2010 - 2013				Total	2014 - 2017				Total	Grand Total
	IS	Med	Both	Other		IS	Med	Both	Other		IS	Med	Both	Other		
Total	**1**	**8**	**2**	**1**	***12***	**3**	**20**	**7**	**1**	***31***	**11**	**72**	**28**	**5**	***116***	**159**
Methodology																
Randomized Controlled Trial (RCT)	0	3	0	0	***3***	0	16	4	0	***20***	0	34	9	2	***45***	**68**
Non RCT intervention/experiment	0	5	1	0	***6***	1	3	0	1	***5***	1	20	8	2	***31***	**42**
Qualitative & ethnography	0	0	0	1	***1***	2	0	2	0	***4***	5	7	4	1	***17***	**21**
Field usability test	1	0	1	0	***2***	0	1	1	0	***2***	5	5	5	0	***15***	**19**
Survey	0	0	0	0	***0***	0	0	0	0	***0***	0	4	1	0	***5***	**5**
Retrospective analysis	0	0	0	0	***0***	0	0	0	0	***0***	0	2	1	0	***3***	**3**
Research Objectives*																
Intervention design/evaluation	0	7	0	0	***7***	0	14	3	1	***18***	0	40	10	4	***54***	**79**
ITSM development and assessment	1	0	1	0	***2***	0	1	3	0	***4***	6	15	8	0	***29***	**35**
Use experience & perceptions	0	0	1	1	***2***	2	5	1	0	***8***	9	6	8	0	***23***	**33**
Chronic care outcome explanation	0	1	0	0	***1***	0	4	0	0	***4***	0	7	2	0	***9***	**14**
Other	0	0	0	0	***0***	0	1	0	0	***1***	1	0	1	1	***3***	**4**
IT Types*																
Mobile/tablet app	1	0	0	0	***1***	0	4	5	0	***9***	10	26	12	0	***48***	**58**
Website	0	0	1	1	***2***	0	4	3	1	***8***	1	19	10	2	***32***	**42**
Medical device	0	2	0	0	***2***	1	1	1	1	***4***	0	19	9	2	***30***	**36**
Smart wearable	0	0	0	0	***0***	0	1	0	0	***1***	3	9	3	0	***15***	**16**

Table 1.2 Profile of the Studies by Discipline and Year

	2006 - 2009				*Total*	2010 - 2013				*Total*	2014 - 2017				*Total*	Grand Total
	IS	Med	Both	Other		IS	Med	Both	Other		IS	Med	Both	Other		
Pedometer	0	1	0	0	***1***	0	1	2	0	***3***	0	9	0	1	***10***	**14**
IVR	0	2	0	0	***2***	0	0	1	0	***1***	0	2	1	1	***4***	**7**
PDA	0	3	1	0	***4***	0	7	0	0	***7***	0	1	0	0	***1***	**12**
PC software	0	2	0	0	***2***	0	3	0	0	***3***	1	1	0	0	***2***	**7**
Other	0	0	0	0	***0***	1	1	0	0	***2***	5	7	3	0	***15***	**17**
Chronic Condition Types*																
Overweight/obese	1	2	0	0	***3***	0	14	3	0	***17***	1	26	6	0	***33***	**53**
Diabetes	0	1	0	0	***1***	0	1	2	1	***4***	3	21	6	2	***32***	**37**
Psychiatric	0	1	0	0	***1***	0	0	0	0	***0***	3	8	4	0	***15***	**16**
Cardiac	0	1	0	0	***1***	0	0	0	0	***0***	1	0	3	1	***5***	**6**
Cancer	0	0	0	0	***0***	0	0	0	0	***0***	1	4	1	0	***6***	**6**
Nerve-related	0	0	0	0	***0***	0	0	0	0	***0***	2	1	0	1	***4***	**4**
HIV	0	1	0	0	***1***	0	0	0	0	***0***	0	3	0	0	***3***	**4**
Hypertension	0	0	0	0	***0***	1	1	0	0	***2***	1	0	1	1	***3***	**5**
Other	0	2	2	1	***5***	1	4	3	0	***8***	5	4	9	0	***18***	**31**

*One study can have multiple research objectives, examine multiple chronic conditions and use multiple IT.

1.6 Results

The 159 studies in our sample are synthesized using the affordance actualization framework, extended to include a high-level summary of the key concepts and relationships which emerged (see Figure 1.3). Through the analysis of a concept matrix (Webster and Watson 2002) and iterative discussions between the researchers, four themes emerged which represent the research in this area: ITSM affordances and related IT functionalities, effects on ITSM use and experience, effects on chronic care goal achievement, and the role of intermediate outcomes. The first theme on ITSM affordances is descriptive in nature, themes two and three are DV-centric, and theme four is concept-centric.

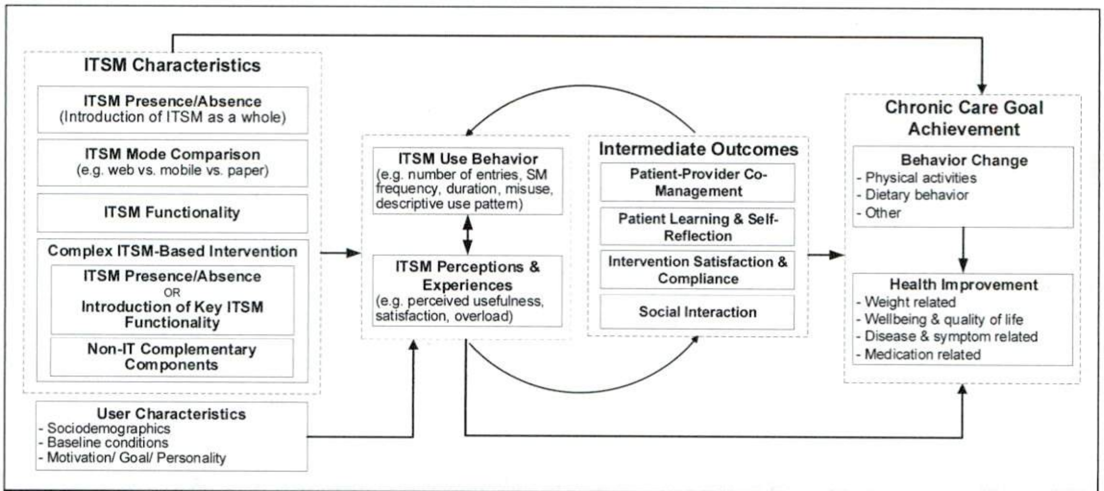


Figure 1.3 High-level Synthesis of Research on ITSM for Chronic Care

1.6.1 Theme 1- Identification of ITSM Affordances and Related IT Functionalities

Theme 1 is descriptive in nature and attempts to capture and categorize the IT functionalities – and the affordances they enable - that are present in the ITSM devices used in the studies. Multiple ITSM devices which provide a range of IT functionalities are represented in the literature, but these affordances and functionalities were seldom directly investigated in the studies (*ITSM characteristics* is used in Figure 1.3 as a more general term to indicate the broad range of ways in which ITSM was investigated in the studies).[5] Although seldom directly examined, it is still worth untangling the functionalities and affordances of ITSM: while the technology itself is rapidly advancing, the associated affordances are likely to evolve more slowly. By understanding ITSM functionalities, we can identify key affordances, understand how these affordances are currently delivered, and reflect on how they may be delivered in the future.

We first coded IT functionalities present in the studies and, drawing on Li et al.'s (2010) model, identified four categories of ITSM affordances. Table A4 in Appendix A presents the ITSM affordances, their associated IT functionalities, and the studies with ITSM that included these functionalities.

1.6.1.1 Preparation Affordance

The importance of training and motivating the actors who engage in chronic care is an essential step (Bodenheimer et al. 2002). The general expectation is that if users are well trained and highly motivated, they are more likely to have sustained engagement which produces better outcomes (Standage et al. 2008; Suh 2018). We have identified two key IT functionalities that support user preparation. First, IT is a low-cost medium to deliver educational content regarding the use of ITSM devices, the knowledge of diseases, the benefit of treatments, and self-management techniques (e.g. Cadmus-Bertram et al. 2013; Dorsch et al. 2015; Or and Tao 2016). These educational materials are often provided as web pages or video clips. Second, many ITSM technologies provide goal setting functionalities that either recommend or prescribe a goal and/or allow users to set or adjust their own goals (e.g. Painter et al. 2017). Some systems are more flexible, allowing users to create detailed action plans, so that they can track goals over time (e.g. Dennison et al. 2014). Although goal setting is very common for chronic care, it is traditionally prescribed by physicians orally or in a written document. More recent studies tend to leverage IT

[5] The affordance-actualization lens views affordances as possibilities for action arising from both IT functionalities as well as actors' needs and goals. In our sample, research often states the chronic conditions under study, with the implied intervention goal being to reduce or manage the symptoms related to the chronic condition. However, an actor's specific internal goals and needs are not explicitly investigated. Thus, our synthesis focuses almost exclusively on IT functionalities.

functions to assign the goal, update the goal based on self-tracked progress and keep track of the goal changes digitally.

1.6.1.2 Data Collection Affordance

Since recording is the core activity of ITSM, all ITSM devices should provide functionalities to support data collection and/or data entry, either fully automated or requiring a certain level of human effort. Many ITSM systems have a data entry interface that enables user-initiated entry of SM data. For example, the user may have to manually measure certain SM data of interest (e.g. weighing dietary intake, self-assessing mood), and then use the ITSM interface to record it in the system. The data entry user interface may involve different levels of flexibility such as guided response (e.g. structure daily questionnaire, Tsanas et al. 2016) or open entry (e.g. journaling, Hales et al. 2017).

Due to the development of sensor technologies, various activities, positions, proximities and body conditions can be detected automatically without active human effort, making it one of the most significant advantages of ITSM over traditional paper- or memory-based approaches (Lupton 2014). The most widely-used devices with auto data capture functionalities are wearables such as fitness trackers and pedometers. While wearables such as pedometers have been used for many years, more recent studies employ smart wearables which support multiple affordances in an integrated manner (e.g. automatic data capture, interactive data display, goal updating, pushed tips). There is increasing attention in recent years regarding insideables, such as under-skin continuous glucose monitoring devices. Although the technology has been available for over 15 years, major health insurances companies in Europe and North American have only started to cover the devices in recent years (Heinemann and DeVries 2016). The increasing availability and popularity may foster future studies on the use and implementation of these new ITSM technologies. In recent years, ITSM with automated data capture is becoming more widely used and more frequently studied. Despite the convenience, fully automated data capture is not possible or appropriate for all SM tasks. Many diary SM tasks still require significant manual measurement (e.g., weighing the food) with user entry. For some tasks where automated data capture is possible, a data entry interface may still be required where users can override the automatically captured data or correct erroneous readings (Selvan et al. 2017).

1.6.1.3 User Reflection and Action Affordance

Three key IT functionalities emerged in our review related to data display, push messages and gamification, which should contribute to *user reflection and informed action*, a key stage of SM (Li et al. 2010). According to the definition of SM, if patients are not allowed to view and/or use their own data, it is surveillance rather than SM. With data displays being increasingly digital, there is a trend towards increased transparency provided to patients rather than physicians controlling the flow of data (Cade 2017; Piras and Miele 2017). Three levels of graphical, numerical, or text feedback of the SM results have been found: (1) raw data is presented in graphs, tables or text (e.g., readings from a glucometer); (2) aggregations of the data are presented, such as total number, average and calculated indices (e.g. energy expenditure based on activity energy consumption and diet energy intake, Allen et al. 2013); and (3) evaluative information is provided that relates the data to a target, goal or threshold. This last type is commonly presented using colored traffic light systems (e.g. blood glucose levels, Greenwood et al. 2015), progress bars to show performance as compared to the desired goal (Carels et al. 2017), or textual messages that provide personalized assessments (Wolin et al. 2015). Usually, these data displays are the result of requests from the user to access this information (a "pull" type of communication). The data display influences users' ability to make sense of their data, thus supporting user reflection and action.

We have also found some devices provide "push" communication, where feedback is sent to users in the form of prompts, alarms, reminders, and push notifications (e.g. Ambeba et al. 2015; Chambliss et al. 2011). Push messages are an important tool in persuasive computing (Oinas-Kukkonen and Harjumaa 2009) and should have positive effects on user awareness and engagement in ITSM. Two types of push messages emerged in the reviewed literature: pre-set and data-driven. Pre-set push messages are usually time-based, with users or algorithms setting alarms or reminders for specific SM tasks (Swendeman et al. 2015). Data-driven messages are usually triggered by events related to users' performance or progress (e.g. prompts after a prolonged period of inactivity, Biddle et al. 2017).

Gamification – defined as using game design elements in non-game contexts (Deterding 2015) – is one often used approach in persuasive computing that is considered an effective way to enhance user experience and promote performance in many different contexts (Mekler et al. 2017), Thus, it should positively influence user motivation and engagement in ITSM and is a functionality to support reflection and action. Some ITSM represent the SM task in a gamified way (e.g. simulations and challenges), or the data are displayed with gamified elements such as reward points and leader boards (e.g., Hales et al. 2017; Hostler et al. 2017). The design and use of gamification in ITSM is quite rare in our sample until 2017, when 13 studies were published that often involve design science research focusing on specific feedback or incentive mechanisms.

1.6.1.4 Social Connection Affordances

Support for social connections is not included in Li et al.'s (2010) five stages; however, its importance is recognized by chronic care practitioners (Wagner et al. 2001). Two categories of IT functionalities that enable such affordances emerged from the literature. The first category supports patient-provider connections by providing contact information of the healthcare team or allowing synchronous or asynchronous online communication (e.g. Greenwood et al. 2015; Iljaz et al. 2017; Webber et al. 2010). The second category supports peer-to-peer interaction, where a virtual space is created for peers (including other patients and any non-provider trusted entities such as friends, acquaintances and caregivers) to exchange information and influence each other (e.g. Cadmus-Bertram et al. 2013; Mummah et al. 2017). Some virtual spaces function using private groups of existing external social networks or online communities (e.g. Cadmus-Bertram et al. 2013; Carter et al. 2013; Partridge et al. 2016), while other recent mobile-based trackers include embedded social functions for within-app communities or links for sharing to mainstream social media platforms (e.g. Eikey et al. 2017; Hales et al. 2017).

1.6.1.5 ITSM Affordances and Bundles by Disease type

SM requirements should align with chronic care goals, and – in practice – the choice of ITSM device largely depends on disease type. Thus, we further analyzed the presence of ITSM functionalities and the associated affordances by disease type to examine which combinations are being studied (see Table 1.3).

It is not surprising that all studies have at least one type of data collection method since self-recording is a fundamental task for SM. SM tasks for obesity usually involve physical activity and dietary intake, so both automatic data capture (for exercise) and manual input (for dietary intake) are very common. Diabetes self-management usually involves blood glucose monitoring, thus auto-capture by medical device is more common. Whereas older approaches usually involve two devices (glucometer plus another database-type application with manual data transmission between the two, Roblin 2011; Sevick 2008), recent research usually investigates glucometers with mobile or web applications that allow automatic data transmission or sync functions (Garg et al. 2017; Sieber et al. 2017). Moreover, with the increasing popularity of under-skin sensors for continuous blood glucose monitoring, patients no longer need to worry about data capture and entry (Polonsky et al. 2017). Managing psychiatric conditions usually involves answering questionnaires regarding mood or psychological issues; thus, guided manual entry is dominant.

Data display is the second most common functionality. However, the majority of studies only display raw data or descriptive information after simple aggregation (e.g. Aguiar et al. 2017). This is more apparent for studies using pedometers and older-style glucometers that only support basic data collection and display. Push messages and gamification are two functionalities that have started to receive attention in recent years. Most of them are used for exercise and diet-related SM tasks, both related to obesity, perhaps because the system can more easily generate meaningful time-based push messages (e.g., reminders to enter meal information three times per day, reminders for physical activity after long periods of sedentary behavior). Data-driven push messages are rare (for an exception, see Coppini et al. 2017), requiring more personalized messages and data analytics effort.

While by definition, SM requires some level of data collection and reflection (whether IT-based or not), preparation and social connection seem to be optional affordances. Recent years have witnessed an increasing number of studies with IT-supported education, goal setting, patient-provider connection and peer-to-peer interaction. Most of the preparation affordances are present in obesity self-management through mobile or web applications, making it easy for digital materials to be presented and adjusted during the longitudinal intervention. Although goal prescription is a common strategy for most healthcare interventions, IT-based goal setting and adjustment functionalities mostly appear in fitness tracking and diet applications and are therefore most common for obesity SM. The patient-provider connection function is increasing, even though only part of the interaction may be directly supported by the SM device (e.g. physician receives system notification then sends email to the patient). Instant communication with healthcare providers within the same device is limited. The incorporation of this kind of functionality may be largely constrained by clinical practices, such as healthcare providers not constantly monitoring patients' conditions, or the clinical infrastructure not being able to support this particular channel of communication. Similarly, peer-to-peer interaction via social components in ITSM appears more common for exercise and diet SM, which is partially due to the fact that exercise and diet information is less sensitive, and peer comparison has been shown to be a useful approach to promote exercise and weight loss by controlling for diet (Finnerty et al. 2010; Luszczynska et al. 2004; Mueller et al. 2010; Thompson et al. 2006).

Table 1.3 Presence of ITSM Affordances by Chronic Disease Type

Affordance	Functionalities	Chronic Disease Type							
		Obesity	Diabetes	Psychiatric	Cardiac	Cancer	Nerve-related	HIV	Hyper-tension
	Total	*53*	*37*	*16*	*6*	*6*	*4*	*4*	*5*
Preparation	Education	16	7	2	2	2	1	1	2
	Goal	13	1	1	0	0	0	0	0
Data collection	Data entry	45	15	14	2	2	2	4	2
	Auto capture	21	30	2	4	3	3	0	2
Reflection and Action	Data display	40	32	12	5	3	3	3	4

Table 1.3 Presence of ITSM Affordances by Chronic Disease Type									
		Chronic Disease Type							
Affordance	Functionalities	Obesity	Diabetes	Psychiatric	Cardiac	Cancer	Nerve-related	HIV	Hyper-tension
	Total	*53*	*37*	*16*	*6*	*6*	*4*	*4*	*5*
	Push message	24	9	5	1	2	1	2	3
	Gamification	4	2	3	0	0	1	0	0
Social Connection	Patient-provider connection	5	8	3	0	2	1	0	0
	Peer-to-peer interaction	14	1	1	0	0	0	0	0

It should be noted that the ITSM process is longitudinal, flexible, and iterative, and for the use of a specific ITSM device multiple affordances may be bundled to provide different action possibilities (i.e. termed *affordance bundles* in Strong et al. 2014). For example, if personalized education can be delivered at the right time (e.g. when the indicator goes above a certain threshold) with push messages, patients may be more capable of understanding the data and take appropriate action as necessary, such as initiating a conversation with physicians (e.g. Caballero-Ruiz et al. 2017; Velardo et al. 2017). Thus, the effectiveness of ITSM may depend on bundles of affordances and their actualization, rather than an isolated functionality and its associated affordance.

1.6.1.6 Theme 1 Discussion and Future Directions

A summary of key results and future research directions for theme 1 are shown in Table 1.4. The IT functionalities synthesized above afford users the ability to perform a multitude of key SM steps (Li et al. 2010) by allowing users to complete these tasks more efficiently (e.g., smart fitness trackers that automate data collection and provide real-time data display, Abrantes et al. 2017) or by allowing users to complete new tasks that were not easily accomplished by pure human effort (e.g. continuous blood glucose capture, Polonsky et al. 2017). Ideally, ITSM affordances will enable individuals to perform SM tasks, increase their SM motivation and preparedness, and better incorporate SM results into their chronic care practices.

It should be noted that some functionalities and affordances are more basic and fundamental (e.g. data collection and data display), whereas others are complementary to provide added value and may be delivered in a separate device (e.g. using social media in addition to the wearables for peer support). There is a general trend of ITSM devices becoming increasingly multifunctional and interconnected. As technologies advance, how these affordances can be delivered (i.e. IT functionalities) may change dramatically. By linking IT functionalities into higher-level ITSM affordances, we bring to light the key capabilities that an ITSM system could deliver and describe current IT functionalities used to provide these capabilities. However, despite the effort that has been made to design the new applications and system prototypes (e.g. Cai et al. 2017; Coppini et al. 2017), less is known regarding whether the emerging functionalities universally improve delivery of the designated affordances and ITSM efficacy. Moreover, as new medical technologies become increasingly mature (e.g. insideables, artificial intelligence applications), these devices in various forms and formats may bring new affordances or make existing affordances obsolete, which may bring interesting synergistic effects that we cannot foresee with the current generation of technology.

Accordingly, we propose three promising areas for future research. First, future research can focus on investigating how to better deliver the four ITSM affordances that have been identified. For example, as a basic function, data display and feedback functionalities are currently built around simple descriptive statistics, which can be trivial given the complexity of patients' experience living with chronic conditions. With the advancement of data analysis techniques such as machine learning and natural language processing, it is possible to provide personalized and explanatory feedback that can reveal the causal linkages behind the patient's living trajectory (Piwek et al. 2016). Studies on wearables and mobile-based self-tracking for general populations have made various attempts to design the system with better usability and entertainment (e.g. Leinonen et al. 2017; Liang et al. 2017; Tay et al. 2017). ITSM for chronic care can incorporate these research achievements and develop more context-specific solutions.

Second, there are many emerging technologies that are not currently widely used in the market but are promising and may have already shown profound impacts in other areas.

For example, artificial intelligence (AI) has received tremendous attention in many areas, and in healthcare, AI also has potential for a wide range of applications, from diagnostics to operations (e.g., fraud detection, virtual nursing, medical error reduction, automated diagnosis, Kalis et al. 2018). These new trends bring advanced functionalities and capabilities that may not only improve healthcare efficiency but also shift how patients and providers perform ITSM. For example, AI may be particularly useful for chronic diseases where the link between cause and effect for a particular patient (e.g., triggers for a patient's migraines) are not always evident. For these chronic diseases, using ITSM with advanced AI functionalities that can discern the complex patterns between triggers and symptom onset for one specific patient may help that patient predict (and ultimately, avoid) these triggers. Future research should explore new technological developments, the impacts of their associated functionalities and if these technologies modify how affordances are delivered in ITSM.

Finally, we propose that future research should highlight the context of new applications and investigate the conditions under which those advanced functionalities and add-on affordances engender positive effects versus conditions that may yield negative consequences. Existing studies have reported some functional barriers in adoption and use of ITSM (e.g., Chung et al. 2015) and potential negative side-effects such as information overload (Mathieu-Fritz et al. 2017), and gamification elements acting as distractions from the main SM tasks (Sage et al. 2017). Research on social support is nascent and further research is needed to examine the role of social connections. For example, future ITSM may include functionality which shares data with caregivers at opportune moments, especially in cases where the patient can't help themselves or the nature of the disease makes it unlikely that they will seek help (e.g., a system that notifies the caregivers when patient SM data shows an increase in markers related to the disease). Future research should investigate both positive and negative aspects of advanced functionalities and add-on affordances.

Table 1.4 Summary of Theme 1
What is known

• Current ITSM has four key affordances:
  ○ Preparation.
  ○ Data collection
  ○ User reflection and action.
  ○ Social connection.
• Some affordances are more fundamental to all ITSM (e.g. data collection, user reflection and action), while others may enable ITSM (preparation) or act as optional add-ons (social connection).

What is unknown and suggestions for future research

• If emerging IT functionalities universally improve delivery of these affordances.
• Whether new and emerging IT functionalities generate new affordances or make existing affordances obsolete.
  ➢ Determine how to better design IT functionalities to improve delivery of these ITSM affordances.
  ➢ Investigate how future and emerging IT functionalities change how these affordances are delivered.
  ➢ Explore concept of new affordances and affordance obsolescence.
• Whether optional add-on affordances add value to or provide a distraction from SM.
  ➢ Examine under which conditions add-on functionalities and affordances engender positive effects and under which conditions they act more as distractions from the main SM tasks.

1.6.2 Theme 2- Effects on ITSM Use and User Experience

Theme 2 examines two closely related components of the affordance actualization process: ITSM system use (shortened to ITSM use) and user experience. Given that intervention compliance is especially important for chronic care (Hamine et al. 2015), both sufficient usage of ITSM devices and positive patients' experiences should be essential. Seventy-two studies report evidence regarding patients' use of ITSM and their perceptions. These studies employ various research methods (e.g. experiment, survey, secondary data analysis, field study, interview, and focus group). The majority focus on ITSM use frequency, with a small number examining use duration, misuse, appropriation, use patterns, etc. ITSM use frequency is usually measured by number of SM entries, number of days with logins, frequency of feature use, or frequency of SM website visit. The key constructs and relationships empirically investigated for theme 2 are shown in Figure 1.4 and Table 1.5.

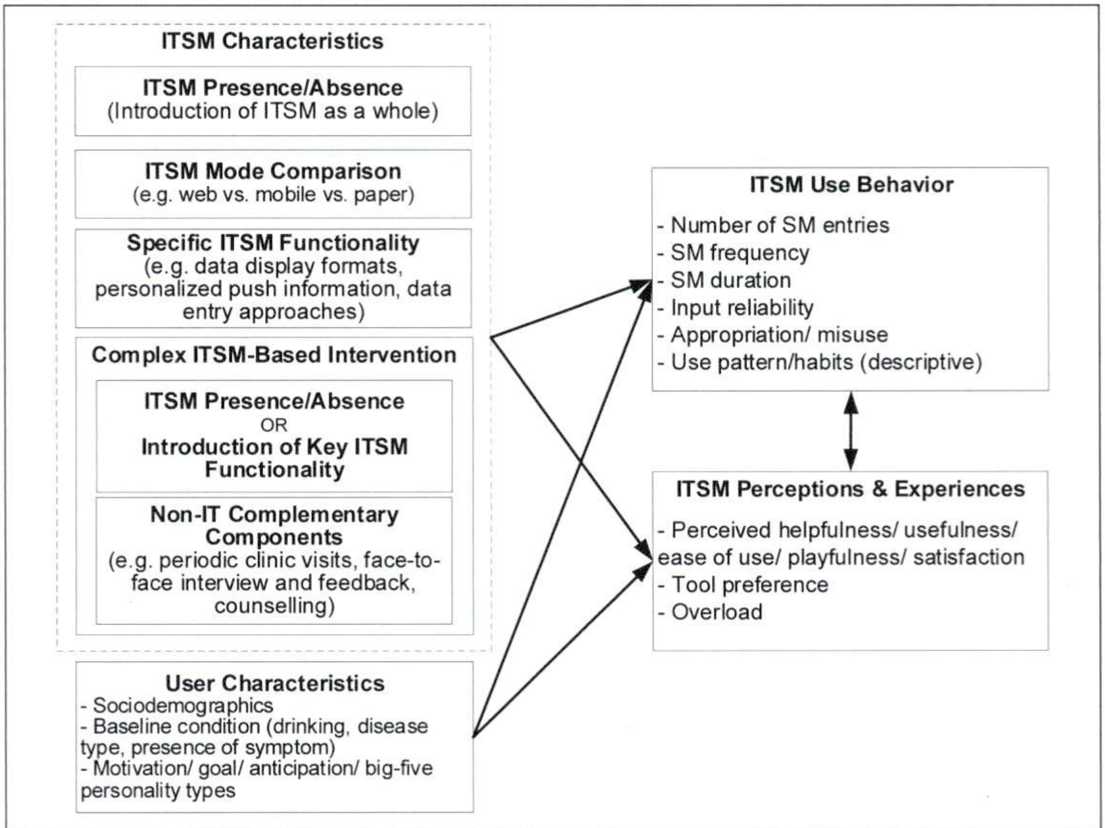


Figure 1.4 Relationships Investigated impacting ITSM Use and User Experience

Five categories emerged from literature as major sources of impacts on ITSM use and experience: ITSM presence, ITSM mode, specific ITSM functionality, complex ITSM healthcare intervention, and user characteristics. Although in theory it is the actualization of specific ITSM affordances enabled by one or multiple IT functionalities that facilitates goal achievement, the extant studies seldom discuss the impacts of specific IT functionality or affordances. Instead, ITSM is usually introduced as a whole system for the entire chronic care procedure or presented as part of a complex intervention in which ITSM is supported by various non-IT healthcare components. Consequently, a typical design of the study includes: (1) introducing and evaluating a complete ITSM program for chronic care (i.e. ITSM presence, Johnston et al. 2009; Velardo et al. 2017); (2) comparing ITSM modes by using different technologies or different SM designs (with or without non-IT components), which in turn provides implications for potential advantages of specific ITSM devices (Or and Tao 2016; Swendeman et al. 2015; Turner-McGrievy et al. 2017); (3) designing new ITSM tools with an emphasis on specific functionalities

(Adams et al. 2017; Sage et al. 2017); and (4) introducing a complex ITSM intervention (ITSM plus non-IT components, Partridge et al. 2016) or comparing multiple complex ITSM interventions (Sevick et al. 2008; Spring et al. 2017). In the following section, we present key constructs and relationships investigated for the impacts of ITSM characteristics on ITSM use and user experience.

When presented with the ITSM as a whole, the studies generally have positive results regarding both SM use (i.e., number of days with entries or acceptable SM rate, Roblin 2011; Tsai et al. 2007) and usability (e.g. usefulness, ease of use and satisfaction, Festersen and Corradini 2014; Timmerman et al. 2016; Gu et al. 2017). Such positive usability evaluations facilitate ITSM frequency (Ma et al. 2013) and engagement (Adams et al. 2017). However, when specific functionalities are selectively assessed, some studies report negative evaluation and use behaviors. For example, during an evaluation of a newly developed system allowing users to freely tag their SM activities, participants report reluctance to use the feature due to difficulties in understanding the new display format (Storni 2011). Similarly, participants negatively evaluate a newly developed app with gamification functions because it requires extra effort, which hinders their engagement (Sage et al. 2017). When patients feel overwhelmed by the system functionalities, they are more likely to misuse the device, develop workarounds, or return to sub-optimal SM practices (Mathieu-Fritz and Guillot 2017).

Several studies formally compare SM frequency and user satisfaction for different modes of ITSM and generally support the superiority of automated SM (e.g. web-based, mobile-based, pedometer) as compared to a paper-based approach. Paper-based SM is rated as inconvenient, embarrassing and less fun (Hutchesson et al. 2015; Fuller et al. 2017), whereas web-based and mobile-based SM exhibits higher use frequency (Or and Tao 2016), is less burdensome (Matthews et al. 2017), has fewer recording errors (Selvan et al. 2017) and is preferred by more users (Hutchesson et al. 2015). However, when comparing web-based and wearable-based SM for diet tracking, Turner-McGrievy et al. (2017) do not find differences in SM frequency. This may be because diet SM data collection – whether web-based or wearable – is not automatic and thus wearables do not provide a significant advantage over other ITSM modes in this context.

The effects of implementing complex ITSM interventions on ITSM use and user experience are highly mixed. Several studies report a "novelty effect" where the ITSM frequency declined rapidly after initial use (Carter et al. 2013; Glasgow et al. 2011; Laing et al. 2015; Stark et al. 2011; Wolin et al. 2015). Such a decline in use may be temporarily averted with IT or face-to-face feedback that is either continuous or novel (i.e., personalized and non-repetitive pushed messages). For example, patients who attend a group counseling session continue ITSM use while use declines for those who missed the session (Sevick et al. 2010). Surprisingly, the post-adoption decline in use is greater among SM web users than interactive voice response system (IVR) users (Wolin et al. 2015). This may imply that while newer modes of ITSM – which are thought to reduce user burden – increase ITSM use initially, the reduced burden and near invisibility of the newer ITSM may make it harder for users to develop sustainable habits. However, due to the complexity of interventions and how they are investigated, it is difficult to know whether novelty effects are due to the IT functionalities or to the non-IT components. Additional studies are needed in this area while newer technologies that automate data capture and entry may largely eliminate human effort in use, they may also yield unintended consequences. For example, ITSM with automatic data capture may not afford users as many opportunities to actively think, learn from, and be aware of their health-related behaviors and conditions.

User characteristics often also influence when and how ITSM is used. In the context of chronic care, user baseline health status is an important factor. A couple of studies examine the impacts of user baseline status on their subsequent use behaviors, including sociodemographic factors such as age, gender and education (Di Bartolo et al. 2017; Sevick et al. 2010; Wolin et al. 2015), current disease condition (such as presence of depression (Steinberg et al. 2014)), family history of diabetes (Cosson et al. 2017), motivation (Webber et al. 2010), conscientiousness (Hales et al. 2017), and SM objectives (maintain normalcy vs. self-stabilization, Mattews et al. 2017). No conclusion can be made with regard to the impact of sociodemographic factors due to the limited number of studies and inconsistent results for each factor. For example, there is conflicting evidence on whether or not older people use ITSM more than younger ones (Berry et al. 2015; Glasgow et al. 2011; Krukowski et al. 2013; Sevick et al. 2010). There are no significant associations between education (Glasgow et al. 2011; Krukowski et al. 2013; Sevick et al. 2010) or marital status (Berry et al. 2015; Krukowski et al. 2013) and ITSM use. These user characteristics are most often examined in healthcare journals, most likely because the research tradition in healthcare recognizes that a patient's sociodemographic information may influence the feasibility and applicability of a given treatment. Thus, user characteristics are treated as predictors in this type of research. However, little theory or explanation is provided for why certain user groups should exhibit more ITSM use than others.

Finally, two studies identify users' conscientiousness and autonomous motivation as predictors of ITSM use (Hales et al. 2017; Webber et al. 2010). In behavior change research, user motivation indicates their willingness and psychological preparedness to adopt a medical intervention or engage in a volitional process (DiClemente et al. 2004; Prochaska and DiClemente 1982). ITSM is largely a volitional process, and a certain level of psychological preparation is necessary to plan the intervention strategy and facilitate long-term user commitment (Biener and Abrams 1991; Daley and Duda 2006; Holt et al. 2010).

Table 1.5 Impacts on ITSM Use and User Experience

Impacts on ITSM use from:	
ITSM presence	Roblin (2011), Storni (2010), Tsai et al. (2007), Tsanas et al. (2016), Welch et al. (2007), Boyd et al. (2017), Isetta et al. (2017), Velardo et al. (2017)
ITSM mode	Raiff and Dallery (2010), Selvan et al. (2017) *[Or and Tao (2016), Turner-McGrievy et al. (2017)]*
ITSM functionalities	Kendall et al. (2015), Murnane et al. (2016) *[Chung et al. (2015), Storni (2014)]*
Complex ITSM intervention	Burke et al. (2012), Cadmus-Bertram et al. (2015), Carter et al. (2013), Cushing et al. (2011), Sevick et al. (2010), Sevick et al. (2008), Turk et al. (2013), Wolin et al. (2015), di Bartolo et al. (2017), Spring et al. (2017) *[Aharonovich et al. (2006), Conroy et al. (2011), Glasgow et al. (2011), Laing et al. (2015), Morgan et al. (2014), Partridge et al. (2016), Stark et al. (2011), Thomas et al. (2015), Turner-McGrievy et al. (2013), Wharton et al. (2014), Aguiar et al. (2017)]*
User characteristics	Aharonovich et al. (2006), Berry et al. (2015), Chung et al. (2015), Hall and Murchie (2014), Webber et al. (2010), Wolin et al. (2015), Cosson et al. (2017), Hales et al. (2017), Matthews et al. (2017a), McDonald et al. (2017)

Table 1.5 Impacts on ITSM Use and User Experience	
	[Glasgow et al. (2011), Karhula et al. (2015), Krukowski et al. (2013), Sevick et al. (2010), Steinberg et al. (2014), di Bartolo et al. (2017), McKnight et al. (2017), Selvan et al. (2017)]
Other influencing factors	Sevick et al. (2010), Turner-McGrievy et al. (2013), Adams et al. (2017), Chen et al. (2017), Isetta et al. (2017), Mathieu-Fritz et al. (2017), Matthews et al. (2017a) *[Ma et al. (2013)]*
Impacts on ITSM perceptions and experiences from:	
ITSM presence	Festersen and Corradini (2014), Caballero-Ruiz et al. (2017), Johnston et al. (2009), Nakano et al. (2011), Roblin (2011), Timmerman et al. (2016), Tsai et al. (2007), Abrantes et al. (2017), Boyd et al. (2017), Coppini et al. (2017), Gu et al. (2017), Gell et al. (2017), Isetta et al. (2017), McDonald et al. (2017), Mouzouras et al. (2017), Olafsdottir et al. (2017), Welch et al. (2007)
ITSM mode	Hutchesson et al. (2015), Raiff and Dallery (2010), Swendeman et al. (2015), Fuller et al. (2017), Matthews et al. (2017a)
ITSM functionalities	Hall and Murchie (2014), Hinnen et al. (2015), Andersen et al. (2017), Cai et al. (2017), Edge et al. (2017) *[Adams et al. (2017), Sage et al. (2017)]*
Complex ITSM intervention	Cadmus-Bertram et al. (2013), Laing et al. (2015), Morgan et al. (2014), Sevick et al. (2008) *[Carter et al. (2013), Ma et al. (2013), Biddle et al. (2017)]*
User characteristics	Ramanathan et al. (2013) *[Hall and Murchie (2014)]*

Note. Studies in italicized brackets have non-supportive or mixed results.

1.6.2.1 Theme 2 Discussion and Future Directions

The key findings of theme 2 – along with directions for future research – are outlined in Table 1.6. Research on ITSM use mostly focuses on SM frequency, which is an essential building block of adherence to chronic care programs. ITSM mode matters with users generally preferring technologies that impose less of a burden.

Only a few theoretical perspectives are used in theme 2 and over two thirds of the studies do not employ any theoretical lens. For those that do, social cognitive theory is the most widely cited. However, it is usually used to inform overall intervention design (e.g., Allen et al. 2013) or to interpret study results (e.g., Cadmus-Bertram et al. 2013, Hales et al. 2017), rather than to support hypotheses specific to theme 2. Theoretical lenses including TAM and UTAUT are also used in a few studies, but only for measurement development (e.g. Laing et al. 2015, Ma et al. 2013). Thus, the extant studies in theme 2 rarely provide theoretical explanations regarding why users adhere to certain ITSM tools or interventions and therefore use them more frequently. However, IS research includes several theories – such as those related to coping (e.g., Beaudry & Pinsonneault 2005, Stein et al. 2015) and habit (e.g., Polites and Karahanna 2013, Wilson et al. 2010) – which could be used in future theme 2 research to better understand why patients with chronic conditions use (or avoid) ITSM (see below for brief illustrations).

In theme 2, ITSM is often presented as a whole, thus we do not know how specific ITSM functionalities fulfill users' needs and expectations, nor how these needs influence use. It is also unclear whether new and more advanced functionalities such as sophisticated data display and gamified SM tasks (e.g. challenges, raising virtual pets) increases ITSM use or are a distraction from the main SM task. Moreover, since the majority of the studies under this theme used complex ITSM interventions (ITSM + non-IT components), it is difficult to tease apart the effects of multiple intervention components. Thus, we do not know whether the mixed results are due to the ITSM or the non-IT components.

A novelty effect is reported in multiple studies, and negative evaluations are reported for newly developed functionalities such as innovative data display. It is possible that users with chronic diseases are a group that may not react to novelty as well as other user groups due to the crucial nature of the SM task. For example, diabetes patients who are already familiar with traditional glucometer measuring approaches may be very hesitant to use new continuous glucose monitoring technologies, in order to avoid any possible errors which may impact their health.

We propose three promising areas for future research. First, although overall evaluation of an ITSM is useful, more research should be conducted to understand how specific ITSM functionalities increase use and enhance user experience. Sources of negative experience and barriers to ITSM use should also be examined, especially as some of the negative experiences emerge in studies of ITSM involving newer functionalities such as gamification (Sage et al. 2017) and more advanced data display formats (Chung et al. 2015, Storni 2014). Theoretical lenses – such as coping theory (Beaudry & Pinsonneault 2005, Stein et al. 2015) that examines IT events, user evaluation, user responses, and nonconforming use patterns – can be useful for understanding negative user experience with ITSM. Second, new forms of ITSM may completely shift how the technology is used, and even how we define ITSM use. With automatic data capture, SM frequency may be of less concern. With insideables that are implanted in the body, SM duration and frequency may not be an issue for intervention adherence. However, issues such as properly applying the tool, data usage and connections with providers are still essential parts of effective ITSM. Future research, drawing on existing IS theories of habit and sporadic use (e.g., Polites and Karahanna 2013, Wilson et al. 2010), can go beyond ITSM use as frequency to more deeply examine different patterns of ITSM use and to examine how emerging IT functionalities influence these use patterns.

Third, the preference of less burdensome ITSM and perception of learning new technologies or functionalities as a burden create an interesting paradox: new ITSM technology may be less burdensome, but it may hinder patients from changing their existing ITSM practices. Introducing new functionality may be difficult for these groups of users where any errors made during the initial learning period with a new ITSM may directly impact their health. Future research should untangle this paradox and examine potential risks and user effort in ITSM for chronic care.

Table 1.6 Summary of theme 2
What is known
• Studies emphasizing new ITSM tool development and usability assessment generally report positive evaluations when the system is introduced as a whole. • ITSM use is measured by use frequency in a majority of the studies. • Certain barriers impede ITSM use: several studies report negative opinions of specific ITSM functions such as unfamiliar data display formats and gamification. • Users generally prefer ITSM technologies that impose less of a burden on them (e.g. automated SM is preferred over paper-based SM). • A “novelty effect” for complex ITSM interventions exists where initial use drops off over time.

Table 1.6 Summary of theme 2

What is unknown and suggestions for future research
• How specific IT functionalities influence ITSM use and perceptions. ➢ Study which specific IT functionalities of ITSM increase use and enhance user perceptions. ➢ Investigate sources of negative user experience and barriers to SM use. ➢ Go beyond ITSM use as frequency to more deeply examine different patterns of ITSM use and to examine how emerging IT functionalities influence these use patterns. • If more advanced IT functionalities (e.g., for automatic data capture) increase ITSM use. ➢ Untangle the observed paradox of how newer automated technologies with less burdensome data collection are preferred for chronic care, as well as whether this technological novelty can also be a barrier for users with chronic diseases. ➢ Investigate the potential roles of effort and risk on ITSM use. • Whether and how user characteristics such as age, education, and lifestyle influence ITSM use and use perceptions. ➢ Empirically examine the impact of a broader range of user characteristics.

1.6.3 Theme 3- Effects on Chronic Care Goal Achievement

Studies examining impacts on chronic care goal achievement are represented by two main pathways: effects of ITSM characteristics (N=79) and impacts of ITSM use and user experiences (N=30). Two frequently examined outcomes required for chronic care goal achievement are behavior change that is related to the task being monitored (e.g., a certain number of steps to take each day), and health improvement that is related to the overarching chronic care goals (e.g., weight loss) (see Figure 1.5).

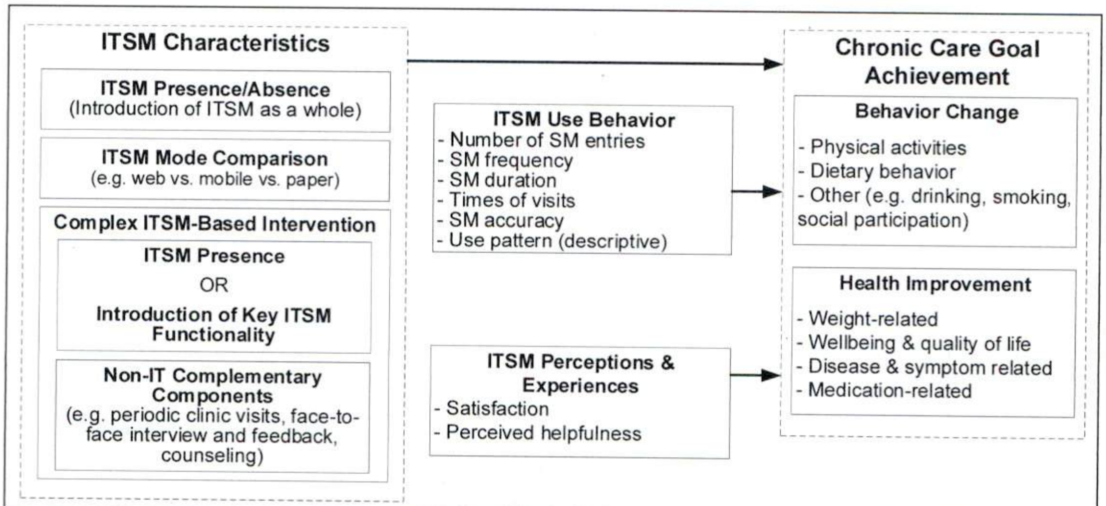


Figure 1.5 Relationships Investigated for Chronic Care Goal Achievement

1.6.3.1 Effects of ITSM Characteristics on Behavior Change

Thirty-six studies examine the direct impacts of ITSM characteristics – whether as ITSM presence, ITSM mode, or complex ITSM interventions – on behavior change. The majority of these studies use randomized controlled trial (RCT) experiments. Two main behavior changes examined in the literature are improving physical activity such as increasing daily steps, physical activity time and reduction of sedentary time (e.g., Cadmus-Bertram et al. 2015; Goto et al. 2014), and improving dietary behavior such as increasing fiber intake, increasing vegetable intake, and balancing calorie consumption (e.g., Ambeba et al. 2015; Jakicic et al. 2017). Some studies examine other behaviors such as the reduction of drinking and smoking (e.g. Swendeman et al. 2015; Aharonovich et al. 2017), both of which are risk factors for psychiatric diseases including depression and bipolar disorder.

The studies suggest that ITSM characteristics generally improve physical activity. Mobile applications and wearables (including pedometers and accelerometers) are the most frequently studied ITSM devices to track steps and physical activity time. ITSM mode comparison studies show that ITSM is better than flexible self-care with periodic counseling (Ruotsalainen et al. 2015), and IT-based SM is more efficient than paper-based SM (Conroy et al. 2011). Type of device (e.g. web-based vs. wearable-based SM) does not seem to influence behavior change. It may be that the presence of ITSM – either using a standalone smart device or multiple connected devices – can perform the simple SM tasks required for physical activity. However, evidence shows that devices providing supportive data display and reflection functions are more effective in improving physical activity (Goto et al. 2014). Several studies report no improvements related to physical activity (see Table 1.7 for a complete list of studies). Potential reasons for these non-supportive results could be using measures that are not directly related to SM tasks, such as sedentary time instead of physical activity time (Biddle et al. 2017; Jakicic et al. 2017) or moderate-to-vigorous PA time instead of PA in general (Abrantes et al. 2017). It should also be noted that the majority of this research employs complex ITSM interventions where the ITSM is combined with multiple non-IT components such as education on self-regulation skills (Morgan et al. 2014), externally prescribed goals (Cadmus-Bertram et al.

2013), or periodic physician reviews of SM results with medical feedback (Nicklas et al. 2014). Thus, it is difficult to clearly attribute the effects of ITSM on physical activity (or lack thereof) to either the non-IT components or the ITSM itself.

ITSM characteristics also have some positive impacts on dietary outcomes, but the results are less conclusive regardless of the design of the intervention or the mode of IT being used. There are two patterns in the non-supportive results: First, most intervention designs do find positive change-from-baseline effects for those with the ITSM, but no significant differences are found between the various ITSM and control groups (Acharya et al. 2011). Thus, the change-from-baseline improvements show the effectiveness of particular ITSM interventions, but no conclusions can be drawn regarding which ITSM intervention design is superior. Second, ITSM characteristics significantly improve general dietary measures such as total calorie consumption, but it does not consistently improve more specific diet indicators such as fiber, sodium, and fat intake (Allen et al. 2013; Jakicic et al. 2016; Schroder 2011; Welch et al. 2013). This highlights the complexity of dietary-related SM tasks, and more research is needed to investigate how to improve specific dietary goals, as certain diseases are more closely linked with specific dietary intakes (e.g. instead of controlling for total calorie consumption, diabetes patients should avoid high carb intake).

Table 1.7 Effects of ITSM Characteristics on Behavior Change	
Impacts on physical activity from:	
ITSM presence/absence	Gell et al. (2017)
ITSM mode	Conroy et al. (2011), Turner-McGrievy et al. (2017) *[Goto et al. (2014), Ruotsalainen et al. (2015)]*
Complex ITSM interventions	Cadmus-Bertram et al. (2015), Cadmus-Bertram et al. (2013), Conroy et al. (2011), Donaldson and Normand (2009), Fukuoka et al. (2011), Izawa et al. (2006), Nicklas et al. (2014), Jakicic et al. (2016), Morgan et al. (2014), Nicklas et al. (2014), Plow and Golding (2017), Vogel et al. (2017), Wang et al. (2012) *[Allen et al. (2013), Abrantes et al. (2017), Biddle et al. (2017), Jospe et al. (2017a), Jones et al. (2014), Laing et al. (2015), Sasai et al. (2017)]*
Impacts on dietary behavior change from:	
ITSM presence/absence	Barakat et al. (2017), Mummah et al. (2017) *[Dowell and Welch (2006)]*

Table 1.7 Effects of ITSM Characteristics on Behavior Change	
ITSM mode	*[Welch et al. (2013), Turner-McGrievy et al. (2017)]*
Complex ITSM interventions	Donaldson and Normand (2009), Fukuoka et al. (2011), Jones et al. (2014), Morgan et al. (2014), Nicklas et al. (2014), Wang et al. (2012), Acharya et al. (2011), Ambeba et al. (2015), Turner-McGrievy et al. (2013), Kempf et al. (2017), Jakicic et al. (2016) *[Allen et al. (2013), Laing et al. (2015), Schroder (2011), Jospe et al. (2017a)]*
Impacts on other behavior changes from:	
ITSM presence/absence	Boyd et al. (2017)
ITSM mode	Swendeman et al. (2015)
Complex ITSM interventions	*[Aharonovich et al. (2006), Abrantes et al. (2017), Aharonovich et al. (2017b)]*

Note. Studies in italicized brackets have non-supportive or mixed results.

1.6.3.2 Effects of ITSM Characteristics on Health Improvement

Sixty-four studies report the direct impacts of ITSM characteristics on health improvement. As with the previous section, a majority of the studies employ complex ITSM interventions. Most of the studies examine weight-related outcomes or disease/symptom improvement, whereas a handful of studies examine quality of life self-assessment and medication change (see Table 1.8).

Table 1.8 Effects of ITSM characteristics on health improvement	
Impacts on weight from:	
ITSM mode	Welch et al. (2013), Turner-McGrievy et al. (2017) *[Ruotsalainen et al. (2015)]*
Complex ITSM intervention	Acharya et al. (2011), Burke et al. (2012), Cadmus-Bertram et al. (2013), Carter et al. (2013), Chambliss et al. (2011), Dennison et al. (2014), Karhula et al. (2015), Morgan et al. (2014), Nicklas et al. (2014), Schroder (2011), Shuger et al. (2011), Sidhu et al. (2016), Steinberg et al. (2013), Thomas et al. (2015), Turk et al. (2013), Turner-McGrievy et al. (2013), Wharton et al. (2014), Aguiar et al. (2017), Carels et al. (2017), Kempf et al. (2017), Jakicic et al. (2016), Moho Shaiful et al. (2017), Munster-Segev et al. (2017), Nishimura et al. (2017) *[Allen et al. (2013), Jones et al. (2014), Laing et al. (2015), Wang et al. (2012), Abrantes et al. (2017), Jospe et al. (2017a), Spring et al. (2017)]*
Impacts on disease/symptoms from:	
ITSM presence	Dietrich et al. (2017), Downing et al. (2017), Sieber et al. (2017), Gell et al. (2017) *[Nørregaard et al. (2014), Umapathy et al. (2015)]*

Table 1.8 Effects of ITSM characteristics on health improvement	
ITSM mode	Goto et al. (2014) *[Or and Tao (2016), Goffinet et al. (2017)]*
Complex ITSM intervention	Chambliss et al. (2011), Karhula et al. (2015), Naylor et al. (2008), di Bartolo et al. (2017), Haak et al. (2017), Hansen et al. (2017), Iljaz et al. (2017), Ji et al. (2017), Kempf et al. (2017), Jakicic et al. (2016), Mantani et al. (2017), Moho Shaiful et al. (2017), Munster-Segev et al. (2017), Nishimura et al. (2017), Sasai et al. (2017), Steinberg et al. (2017) *[Abrantes et al. (2017), Garg et al. (2017), Jospe et al. (2017a), Simons et al. (2017), Young et al. (2017)]*
Impacts on quality of life from:	
ITSM mode	*[Polonsky et al. (2017)]*
Complex ITSM intervention	*[Karhula et al. (2015), Ryan et al. (2012), di Bartolo et al. (2017), Young et al. (2017)]*
Impacts on medication from:	
ITSM presence	*[Dietrich et al. (2017)]*
Complex ITSM intervention	Naylor et al. (2008), Aharonovich et al. (2017a), Aharonovich et al. (2017b), Kempf et al. (2017) *[Pedersen et al. (2012), Plow and Golding (2017)]*

Note. Studies in italicized brackets have non-supportive or mixed results.

Weight management is usually associated with physical activity and/or diet SM, and ITSM demonstrates consistent improvement in weight management when compared to paper SM (Morgan et al. 2014; Wang et al. 2012) or presented together with non-IT components such as counseling and feedback (Cadmus-Bertram et al. 2013; Kempf et al. 2017). However, there is no obvious trend regarding which device performs better, or which part of the complex ITSM intervention is more essential. Several studies report non-significant between-group effects (Jones et al. 2014; Shaiful et al. 2017), and when the weight outcome is measured by BMI or body composition, the results become increasingly inconsistent, even for the change-from-baseline effects (Jones et al. 2014; Ruotsalainen et al. 2015; Jakicic et al. 2016).

For disease and symptom-related improvement, the number of studies for each type of symptom is small—for example, blood pressure in Laing et al. (2015), joint function in Umapathy et al. (2015), depression in Faurholt-Jepsen et al. (2015) and asthma symptoms in Ryan et al. (2012). All of these studies yield mixed results, so no general conclusions can be made. A relatively frequent health outcome that has been reported is HbA1c (average blood glucose sugar levels) for diabetes studies. However, the results are generally non-supportive regardless of SM approach (tablet vs. paper plus glucometer, Or and Tao 2016) or complex ITSM intervention used (e.g. with or without feedback, education or counseling, Greenwood et al. 2015; Young et al. 2017).

Likewise, although positive results are reported regarding medication change and certain aspects of self-rated quality of life, there are not enough studies to reach any general conclusions. No study finds significant improvement for all aspects of quality of life, perhaps because the measurement scale used (usually SF-36 questionnaire, Brazier et al. 1992) includes various aspects of life which do not directly relate to SM goals.

1.6.3.3 Effects of ITSM Affordance Bundles on Chronic Care Goal Achievement

In the theme 3 studies, ITSM is often examined as whole systems which contain various bundles of ITSM functionalities. Although studies did not explicitly investigate the effects of specific affordances, we organize the ITSM affordance bundles from theme 3 in order to see if any patterns emerge (see Appendix Table A2).

In general, automatic data capture with data display exhibits more consistent supportive results across all types of goal achievement than interventions with manual data entry. When push messages are employed, the intervention seems more effective regardless of the data collection approach, except for symptom and medication-related outcomes. Surprisingly, ITSM that supports goal setting and manual data does not exhibit improvement for behavior change and health outcomes. One possible explanation is that patients might have a stronger tendency to misreport the data or adjust goals in order to make the SM results look good and match goals even when no real progress is being made. Since the studies do not report how patients set their goals and perform SM data entry, more investigation is needed to understand under what conditions goal setting, along with other functionalities, is beneficial.

1.6.3.4 Effects of Non-IT Components on Chronic Care Goal Achievement

Because ITSM is not delivered on its own in many cases but as part of a complex ITSM intervention, we attempt to further explore the effects of the non-IT components. The extant studies rarely employ the controlled factorial designs necessary to untangle and compare the role of specific intervention components (for two notable exceptions, see Allen et al. 2013 and Nishimura et al. 2017), so it is difficult to attribute the success or failure of an intervention to either the ITSM or the non-IT components. Thus, we organize the existing interventions from theme 3 by examining the presence of non-IT components (even if they were not directly examined in the studies) to see if any patterns emerge (see Appendix Table A3).

When no non-IT components are present, the results are highly mixed across all outcomes. Similarly, no patterns emerge regarding the effects of non-IT components on chronic care goal achievement. However, four general types of non-IT components are often present, and we describe these types in order to offer insights for future research. The first type is offline education which involves face-to-face training and counseling regarding the disease and self-management skills. This component is either implemented as a one-time education session before the start of the intervention (e.g. motivation elicitation session, Aharonovich et al. 2017b) or implemented periodically throughout the intervention (e.g. weekly coaching in self-regulatory skills, Cadmus-Bertram et al. 2013). The intensity and adaptability of education content vary across studies, but the majority of interventions adopted a pre-designed approach in which the education material is released to the patients gradually without considering the patients' progress or ITSM use experiences. Employing a factorial experimental design, one study examines the effects of both counseling intensity and mobile SM, finding that groups which had both counseling and mobile SM lost more weight than the groups with only counseling or only mobile SM (Allen et al. 2013). Future research can focus more on education delivery by exploring how to use the data from the ITSM to deliver more personalized training.

The second type is goal-related components, namely goal prescription by physicians and goal adjustment during the intervention (e.g. Aharonovich et al. 2017a). One study did carefully isolate the effects of non-IT components: Nushimura and colleagues (2017) study the effects of structured versus routine SM procedures for patients with diabetes while controlling for all other non-IT components, finding that structured SM improves glycemic control while routine SM improves patients' own SM practices. Physicians may assign a specific target for the SM tasks (e.g. 5% weight loss and at least 150 minutes of physical activity, Allen et al. 2013) or create a detailed action plan (e.g. self-management plan, Karhula et al. 2015). These goals can be process-oriented (e.g. number of SM recordings per day) or outcome-oriented (e.g. calorie expenditure per day). Future research can employ goal theories (e.g. goal-setting theory, Locke and Latham 2002) and explore the effectiveness of various goals under different conditions.

The third type of non-IT component is written and/or oral feedback provided by healthcare providers after periodically reviewing the SM results. While IT-based feedback has the advantage of being provided in real-time, written and/or oral feedback may feel more personal, which may elicit better supervision and reinforcement effects. Future research can further investigate feedback mechanisms and compare different modes of feedback to inform better intervention design.

Lastly, offline social activities are used in several studies, including group exercise sessions (e.g. Shaiful et al. 2017) and group-based competition (e.g. Spring et al. 2017). Compared to online social activities in which IT helps construct virtual groups and peer support infrastructures, offline social activities rely more on the organizer (i.e. interventionists) and the requirement of physical presence may make this component difficult to implement during a longitudinal intervention. However, future researchers could examine whether online social mechanisms (usually easier and cheaper to implement) complement or are substitutes for offline social mechanisms.

Appendix Table A3 also presents various combinations of non-IT components. All combinations exhibit mixed results for most of the outcomes. The only evident pattern that emerges is that non-IT educational components exhibit chronic care goal achievements related to weight, and these effects hold when education is combined with feedback. One reason for this effect may be the complex and non-linear relationship between dietary intake and weight requires extra patient education for the ITSM to be effective. Future research is needed to systematically examine the effects of various combinations of non-IT components. In summary, it is difficult to draw conclusions regarding the impacts of the non-IT components based on existing evidence, since the extant research exhibits many mixed results.

1.6.3.5 Effects of ITSM Use on Chronic Care Goal Achievement

ITSM use, as indicated by SM duration or frequency, has several benefits including improved physical activity level (Conroy et al. 2011) and improved dietary behaviors (Glasgow 2011; Jospe et al. 2017). ITSM misuse patterns such as obsessive use and app manipulation may worsen eating disorders (Eikey et al. 2017). Frequency of ITSM entry and data usage can generally predict weight loss (e.g. Kolodziejczyk et al. 2014; Ma et al. 2013; Painter et al. 2017) and HbA1c change (Irace et al. 2017; Lee et al. 2017; Selvan et al. 2017). However, the impacts on other disease-related outcomes are less clear (insulin and cholesterol level, Williamson et al. 2010; blood pressure, Wolin et al. 2015).

Among the twelve studies that simultaneously examine impacts of ITSM characteristics and ITSM use on goal achievement (see Table 1.9), nine report positive results for both relationships (Burke et al. 2012, Cadmus-Bertram et al. 2013; Conroy et al. 2011; Morgan et al. 2014; Spring et al. 2017; Thomas et al. 2015; Turk et al. 2013; Turner-McGrievy et al. 2013; Turner-McGrievy et al. 2017). One reports that while the complex ITSM intervention has positive effects on weight loss, it does not increase ITSM use (Steinberg et al. 2013). Two others report that even when ITSM use and satisfaction are high, the intervention may not lead to weight loss (Polonsky et al. 2017; Wang et al. 2012). No studies formally test mediation effects of ITSM use on the relationship between ITSM characteristics and chronic care goal achievement. In general, the abundance of studies that examine impacts of ITSM use on chronic care goal achievement provide supportive evidence that it is not only the presence of ITSM, but also the extent of use that drives chronic care achievements.

Table 1.9. Impacts of ITSM Use and User Experience on Chronic Care Goal Achievement	
Impacts on behavior change from:	
ITSM use	Conroy et al. (2011), Glasgow et al. (2011), Turner-McGrievy et al. (2013), Eikey et al. (2017), Jospe et al. (2017b) *[Steinberg et al. (2013)]*
ITSM perceptions & experience	Cadmus-Bertram et al. (2013)
Impacts on health improvement from:	

Table 1.9. Impacts of ITSM Use and User Experience on Chronic Care Goal Achievement	
ITSM use	Berry et al. (2015), Burke et al. (2012), Conroy et al. (2011), Jongen et al. (2015), Kolodziejczyk et al. (2014), Krukowski et al. (2013), Ma et al. (2013), Morgan et al. (2014), Steinberg et al. (2014), Thomas et al. (2015), Turk et al. (2013), Turner-McGrievy et al. (2013), Wang et al. (2012), Webber et al. (2010), Hales et al. (2017), Irace et al. (2017), Jospe et al. (2017b), Lee et al. (2017), Matthews et al. (2017a), Painter et al. (2017), Selvan et al. (2017), Spring et al. (2017), Turner-McGrievy et al. (2017) *[Steinberg et al. (2013), Williamson et al. (2010), Glasgow et al. (2011), Wolin et al. (2015)]*
ITSM perceptions & experience	Cadmus-Bertram et al. (2013) *[Polonsky et al. (2017)]*

Note. Studies in italicized brackets have non-supportive or mixed results.

1.6.3.6 Effects of Behavior Change on Health Improvement

Six studies report the impacts of successful behavior change on improving health outcomes (see Figure 1.6 and Table 1.10). Increased physical activity level in terms of time and daily steps (Conroy et al. 2011; Painter et al. 2017; Turner-McGrievy et al. 2013) and lower fat intake (Kolodziejczyk et al. 2014) are associated with successful weight loss. Impacts of physical activity time on reduction in depression and anxiety are inconsistent, which may be due to the method of measurement (e.g., general exercise vs. moderate-to-vigorous activities, Abrantes et al. 2017). Other studies do not find correlations between physical activity and alcohol use reduction (Abrantes et al. 2017) and between HbA1c improvement and self-rated quality of life (Paula et al. 2017). Since chronic care is a long-term journey, patients may have nested behavioral goals in addition to their overarching health goals. The design of the intervention and related ITSM affordances should help users achieve those more actionable behavioral goals in order to gradually achieve more challenging health goals.

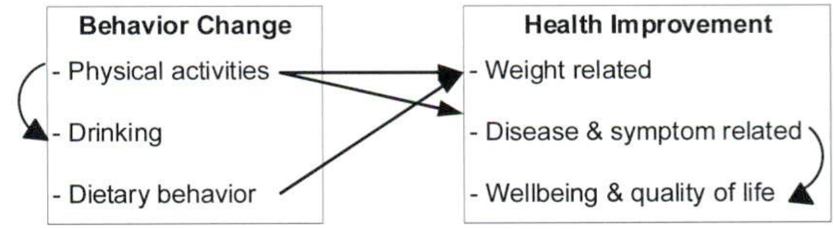


Figure 1.6 Effects of Behavior Change on Health Improvement

Table 1.10 Role of Behavior Change	
Effects of behavior change on:	
Other behavior changes	*[Abrantes et al. (2017)]*
Health improvement	Conroy et al. (2011), Kolodziejczyk et al. (2014), Turner-McGrievy et al. (2013), Painter et al. (2017) *[Abrantes et al. (2017), Paula et al. (2017)]*

Note. Studies in italicized brackets have non-supportive or mixed results.

1.6.3.7 Theme 3 Discussion and Future Directions

To summarize, ITSM research on chronic care goal achievement focuses on behavior change, such as physical activity and diet, and health improvements such as weight management, symptom relief, medication change and self-rated quality of life (see Table 1.11). Most of the studies investigate obesity-related issues, which may explain why physical activity, diet and weight SM have received more attention. Yet the results should have implications for other chronic care contexts, since lack of exercise and being overweight are risk factors linked to many chronic diseases (Fine et al. 2004). The results generally support the positive impacts of ITSM on physical activity, diet and weight reduction, although inconsistent results are reported for between-group differences and some specific measures (e.g. detailed diet indicators, sedentary time, BMI). The results for improving HbA1c in diabetes management and self-rated quality of life are often non-supportive. Overall, these mixed results demonstrate the challenge of using ITSM to change behavior and improve chronic health conditions, although some results do show the positive impacts of ITSM use on behavior change and health improvement.

Similar to theme 2, almost two thirds of the studies in theme 3 do not employ a theoretical lens. For those that do, social cognitive theory is again the most widely used, employed to inform overall intervention or program design (e.g., Allen et al. 2013; Hales et al. 2017) or to interpret study results (e.g., Abrantes et al. 2017; Aguiar et al. 2017). Several studies draw on cognitive behavioral therapy – more of a treatment framework than a theory – to inform intervention design (e.g., Barakat et al. 2017; Mantani et al. 2017). The transtheoretical model of behavioral change (Prochaska and DiClemente 1982) is also used, largely for scale development (e.g., Goto et al. 2014; Izawa et al. 2006). While some theme 3 studies draw on theory and frameworks to inform intervention design, they seldom clearly explain the one-to-one correspondence between theoretical mechanisms and practical intervention components. These theories are mostly used as a background or overarching guidance for the intervention development. For example, social cognitive theory is used to highlight the importance of feedback as a reinforcement mechanism (e.g., Abrantes et al. 2017; Mummah et al. 2017). However, these important mechanisms are not formally tested; instead, they are usually taken for granted as already being part of ITSM. Future theme 3 research should more deeply engage with existing theories (such as social cognitive theory) to examine the impacts of ITSM on goal achievement.

There are several areas of theme 3 that require additional future research. First, chronic care programs can be multifaceted, and complex ITSM interventions are often accompanied by non-IT components. While it may be useful for healthcare practitioners to consider the chronic care intervention as a whole, it is difficult to determine whether the noted improvements (or lack thereof) in behavior and health can be attributed to ITSM alone or to the various non-IT components (e.g., counseling and face-to-face feedback during clinical visits). More effort is needed to untangle the effects of the ITSM intervention components in order to better assess impacts and design more effective ITSM interventions. Such untangling is necessary to understand the synergies between the ITSM and its multiple non-IT components: they may be additive, complementary, or substitutive (Milgrom and Roberts 1995; Samuelson 1974; Titah and Barki 2009). Negative effects may also emerge if the ITSM is too complex and is overwhelming for the patients.

Future research can try to untangle these effects by (1) implementing better controlled experiments with factorial designs, (2) conducting in-depth investigations of specific mechanisms both with and without IT support (e.g. goal-setting mechanisms, feedback mechanisms, social mechanisms), and (3) applying configurational logic (e.g. using qualitative comparative analysis, Schneider and Wagemann 2010) to understand the necessary and/or sufficient components of an effective ITSM design. One example of a better controlled design examines the impacts of counseling by manipulating counseling content and intensity while controlling for mobile SM procedures and feedback components (Allen et al. 2013). Similarly, another study compares structured SM procedures with routine SM procedures while controlling for all the other non-IT components (Nishimura et al. 2017). Both studies found significant between-group differences, suggesting that carefully designing studies to account for the various parts of the complex ITSM interventions may increase the chances of finding clear results.

Second, since ITSM is usually presented as a whole, it is unclear whether the devices with more add-on features – such as an interactive display, real-time communication with physicians, or gamification – yield better outcomes than devices with more basic features. Future research should go beyond simple presence or absence of ITSM to investigate the effects of specific IT functionalities and how they are delivered to impact goal achievement. For example, the incorporation of an incentive system is a new trend in many fitness tracking devices (Hales et al. 2017), and future research should investigate the impacts of incentive design (e.g. process-based vs. event-based incentives, financial vs. virtual incentives). Similarly, with the help of advanced data analysis techniques, future research can compare the effectiveness of different feedback modes (Shin and Biocca 2017), such as comparing feedback format (image vs. textual), timing (event-triggered vs. pre-set), and tone (human-like vs. system-like). Future research in chronic care can focus on how to take advantage of emerging technologies and harness the potential of their functionalities.

Third, more research is needed on health outcomes other than weight loss, as well as behavior outcomes other than physical activity and dietary intake. There are many additional risk factors that are common to multiple chronic diseases and are good candidates for ITSM, such as infections, physiological markers specific to the disease (e.g. metabolome, blood lipids, inflammation) and subclinical symptoms (Tzoulaki et al. 2016). While previous ITSM research was limited by the self-measurement tools commonly available (e.g. pedometer for steps and mobile app for dietary intake), recent technological advances provide more extensive data capture capabilities for the personal collection of various chronic conditions and risk factors (for example, insideables that can track blood glucose levels). A deeper understanding of these technological advancements, how they are changing the delivery of ITSM affordances, and their associated outcomes will allow researchers to investigate a wider range of chronic care issues.

Table 1.11 Summary of Theme 3

What is known

- ITSM presence can help improve physical activity, dietary behavior and weight management.
- IT-based SM is superior to paper-based SM for PA, diet and weight outcomes.
- Complex ITSM interventions exhibit change-from-baseline effects for PA and weight outcomes.
- ITSM presence does not currently improve diabetes management (i.e. does not improve HbA1c).
- ITSM presence does not currently improve self-rated quality of life.
- ITSM use as indicated by SM frequency exhibits positive impacts on PA, dietary and weight management.
- Achieving behavioral goals (e.g. increased physical activity) is beneficial for achieving health goals (e.g. weight reduction)

What is unknown and suggestions for future research

- Whether the noted improvements in behavior and health can be attributed to ITSM alone or to the various components of the complex interventions (e.g., counseling and face-to-face feedback during clinical visits).
    - ➢ Employ more rigorous research designs capable of untangling the effects of complex ITSM interventions.

- Whether the non-significant results of complex ITSM interventions are caused by the ITSM or other components of the complex interventions that are competing sources of influence (e.g., ITSM may improve while gamification may impede health improvements).
- Whether ITSM devices with more add-on features (e.g. interactive display, real-time communication with physicians, gamification) can yield better outcomes than more basic ITSM.
    - ➢ Go beyond ITSM presence/absence and investigate the effects of specific IT functionalities and how they are delivered to impact goal achievement.
    - ➢ Examine ITSM system design with more focus on how to access the potential of more recently available functionalities.

- Whether the impacts of ITSM presence on chronic care goal achievement is mediated by ITSM use and/or intermediate outcomes.
    - ➢ Employ longitudinal designs which capture and analyze the mediating effects of intermediate outcomes.

- Whether ITSM is useful for managing more specific disease symptoms that require complex measurement.
- Which ITSM functionalities are better than others for disease-specific chronic care goals.
    - ➢ Examine ITSM impacts on a wider range of behaviors and health goals.

1.6.4 Theme 4- Intermediate Outcomes of ITSM

The intermediate outcomes of ITSM affordance actualization are the direct results that individuals can achieve due to engaging in ITSM, including psychological or cognitive states induced by the ITSM. The affordance actualization framework suggests that these intermediate outcomes are key mechanisms that help achieve ultimate chronic care goals. The abundance of studies in theme 3 that examine impacts of ITSM characteristics and ITSM use on chronic care goal achievement yield many inconsistent results, implying the existence of these intermediary outcomes. We analyze the ITSM intermediate outcomes in the literature and – through an iterative process of coding, categorization, and research team discussion – find that four categories emerge: patient learning and self-reflection, patient-provider co-management, social interaction with family and peers, and ITSM intervention satisfaction and compliance. In total, fifty-four studies investigate various factors that influence the impacts of, and/or the relationships between, these intermediate outcomes. Although the number of studies for each pair of relationships is small and results are often inconsistent, these studies provide initial evidence regarding the black box between ITSM use and chronic care goal achievement.

1.6.4.1 Intermediate Outcome 1: Patient Learning and Self-Reflection

ITSM can enhance user learning and reflection by delivering education materials and presenting SM data in meaningful ways. As a result, patients can have better comprehension of health problems and SM data so that they can interpret the numbers, find trends, and identify patterns between their SM and chronic conditions (Ayobi et al. 2017; Felipe et al. 2015; Hinnen et al. 2015; Murnane et al. 2016). Three types of intermediate outcomes related to patient learning and self-reflection emerge from the literature: self-understanding, self-efficacy, and health literacy. Figure 1.7 and Table 1.12 present the key constructs and relationships within this theme.

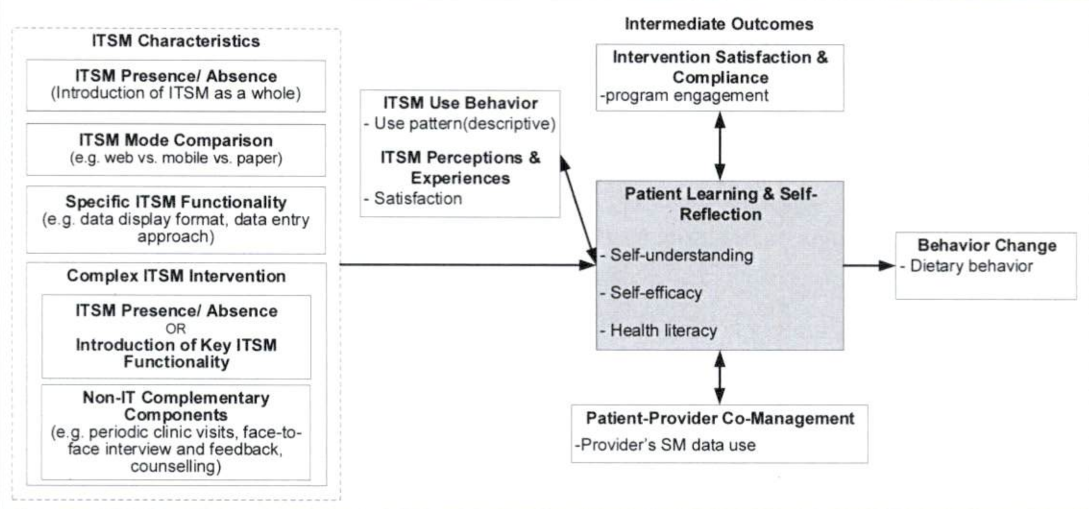


Figure 1.7 Relationships Investigated for Patient Learning and Self-reflection

Table 1.12 Role of Patient Learning and Self-reflection

Impacts on self-understanding from:	
ITSM presence	Ayobi et al. (2017), Bonilla et al. (2015), Murnane et al. (2016), Nørregaard et al. (2014), Tsai et al. (2007), Andersen et al. (2017), McDonald et al. (2017), Velardo et al. (2017) *[Felipe et al. (2015), Gell et al. (2017), Verdezoto and Gronvall (2016)]*
ITSM mode	Swendeman et al. (2015) *[Goffinet et al. (2017)]*
ITSM functionalities	Hinnen et al. (2015), Kendall et al. (2015), Mathieu-Fritz et al. (2017)
Complex ITSM intervention	Aharonovich et al. (2006) *[Cadmus-Bertram et al. (2015), Greenwood et al. (2015), Jones et al. (2014), Webber et al. (2010)]*
ITSM use and experience	Chung et al. (2015), Mathieu-Fritz et al. (2017)
Patient-provider co-management	Chung et al. (2015), Andersen et al. (2017)
Impacts on self-efficacy from:	
ITSM mode	*[Goto et al. (2014), Welch et al. (2013), Polonsky et al. (2017)]*
Complex ITSM intervention	Izawa et al. (2006), Garg et al. (2017), Plow and Golding (2017) *[Greenwood et al. (2015), Laing et al. (2015), Ryan et al. (2012), Rader et al. (2017)]*
ITSM use and experience	Matthews et al. (2017a), Polonsky et al. (2017)
Impacts on health literacy	
ITSM mode	*[Or and Tao (2016)]*

Table 1.12 Role of Patient Learning and Self-reflection	
Complex ITSM intervention	Pedersen et al. (2012) *[Greenwood et al. (2015)]*
Impacts of self-understanding on…	
Behavior change	Bonilla et al. (2015), Kendall et al. (2015)
ITSM use	Eikey et al. (2017)
Intervention compliance	Chung et al. (2015)

Note. Studies in italicized brackets have non-supportive or mixed results.

Patients can use their ITSM data to enhance understanding of their SM results. Patients who habitually use ITSM become increasingly capable of interpreting the results, identifying the correlations, and exploring the causal relations between their daily activities and health conditions (Ayobi et al. 2017; Chung et al. 2015; Felipe et al. 2015; Kendall et al. 2015). Their data interpretation proficiency can be improved through more efficient data display formats (Hinnen et al. 2015) and guidance from clinicians (Anderson et al. 2017; Chung et al. 2015). However, even patients who know how to interpret SM data may not know how to respond in specific health situations and take the right actions (Verdezoto and Gronvall 2016). ITSM can also improve patients' awareness of their self-monitored behavior (e.g. excessive drinking, Aharonovich et al. 2006; dietary intake, Bonilla et al. 2015) and health conditions (e.g. body concern, Ayobi et al. 2017), which helps them foresee the health consequences and prompts preventive and self-regulative actions (Felipe et al. 2015; Murnane et al. 2016; Nørregaard et al. 2014). Two experiments find that ITSM did not improve self-understanding and awareness (Jones et al. 2014; Goffinet et al. 2017). It may be that such awareness is influenced by ITSM design or IT use frequency (e.g. daily vs. bi-weekly SM, Swendeman et al. 2015). The mostly qualitative and descriptive studies that report supportive results provide initial evidence of the impacts on self-awareness, but quantitative analysis is needed to show the actual level of impacts for these relationships.

The impacts on other intermediate outcomes related to patient learning and self-reflection are less consistent. Self-efficacy and motivation level are the two frequently examined concepts in the ITSM literature. Social cognitive theory – which is the theoretical foundation for many ITSM intervention designs – suggests that improving a patient's confidence in his/her ability to self-manage chronic conditions should enhance chronic care goal achievement (Bandura, 1977). A positive correlation between intervention satisfaction and patients' confidence in self-management has also been reported (Polonsky et al. 2017). Yet, evidence shows that it is challenging to improve self-efficacy through ITSM even with carefully designed education sessions (Laing et al. 2015; Rader et al. 2017; Ryan et al. 2012; Welch et al. 2013). One beneficial approach is to help patients improve readiness and motivation, which is reported in several studies (Polonsky et al. 2017; Tsai et al. 2007; Webber et al. 2010). In terms of health literacy, only one study shows ITSM improves patients' disease-related knowledge (Pedersen et al. (2012); however this effect may be due to the other components of the complex ITSM intervention.

Few studies examine the impacts of patient learning and self-reflection on chronic care goal achievement. For example, one study mentions qualitative evidence regarding the beneficial effects of patient awareness and motivation on improved eating habits (Bonilla et al. 2015). More research is needed to investigate the impacts of these intermediate outcomes, as theory would suggest that patients' psychosocial conditions could have profound impacts on chronic care goal achievement (Alderson 1998; Bandura 1998; Deci and Ryan 2008; Walker 2001).

1.6.4.2 Intermediate Outcome 2: Patient-Provider Co-Management of Chronic Conditions

Patient-provider interactions and shared medical decision making is a current trend in chronic care (Bodenheimer et al. 2002; Frantsve and Kerns 2007; Nam et al. 2011). More ITSM devices and interventions are starting to incorporate components that support patient-provider interaction (see Appendix Table A4 for the list of studies with IT-enabled patient-provider connections). Thirteen studies report findings regarding patient-provider co-management as a result of ITSM or the impacts of co-management. Figure 1.8 and Table 1.13 display the key constructs and relationships examined under this theme.

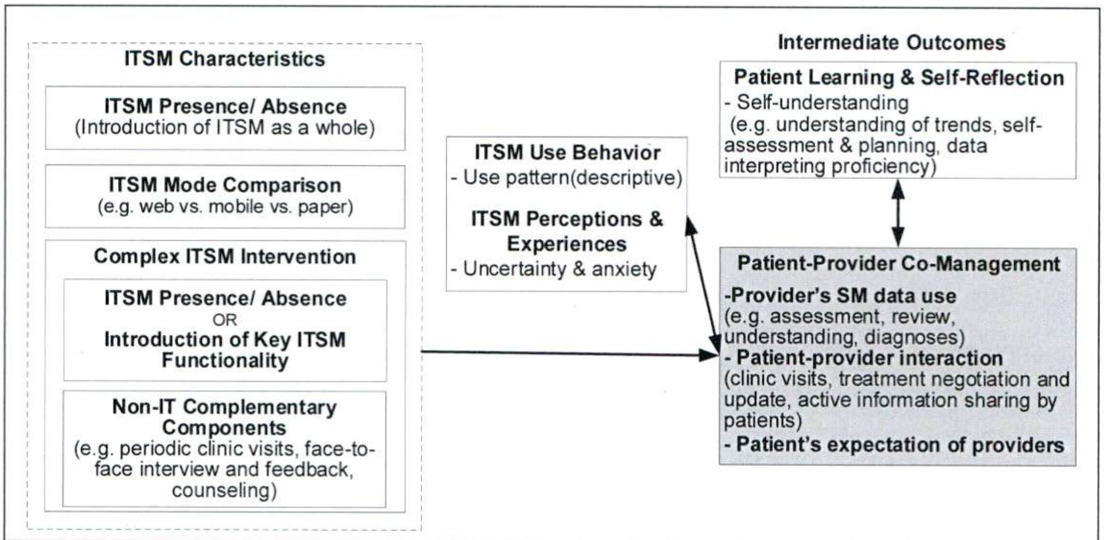


Figure 1.8. Relationships Investigated for Patient-Provider Co-management

Table 1.13 Role of Patient-Provider Co-management

Impacts on Patient-Provider Co-Management from:	
ITSM Presence	Bonilla et al. (2015), Felipe et al. (2015), Andersen et al. (2017), Zhu et al. (2017), Murnane et al. (2016) *[Verdezoto and Gronvall (2016)]*
ITSM Mode	*[Caballero-Ruiz et al. (2017)]*
Complex ITSM Intervention	Nishimura et al. (2017), Rader et al. (2017)
ITSM Use & Experience	Chung et al. (2016), Mentis et al. (2017)
Effects of Patient-Provider Co-Management on:	
ITSM Use & Experience	Andersen et al. (2017), Piras and Miele (2017)
Patient SM Data Use & Self-Reflection	Chung et al. (2015), Andersen et al. (2017)

Note. Studies in italicized brackets have non-supportive or mixed results.

From the providers' perspective, introducing ITSM as a whole or as part of a complex intervention improves the quality and quantity of disease-related information obtained from patients, facilitating assessment, diagnoses and counseling (Bonilla et al. 2015; Murnane et al. 2016). Automatic data-sharing functions reduce the time required by a clinician to integrate the SM records, making it easier for clinicians to review and create personalized treatment plans (Caballero-Ruiz et al. 2017; Zhu et al. 2017). As a result, physicians may change a treatment plan more frequently when ITSM is used (Nishimura et al. 2017).

However, barriers to patient-provider co-management are also reported. First, IT-based tracking is not formally implemented by the majority of clinics (Murnane et al. 2016). Even when ITSM is formally supported, the clinicians may not be fully aware of the system's suggestions or may not trust system-generated information. For example, Caballero-Ruiz et al. (2017) report that the majority of insulin advice provided by the ITSM system was rejected or initially ignored by the medical team. Second, a patient may not be willing to share their self-monitored data with providers (Verdezoto and Gronvall 2016). Patients have their own SM goals and habits which may not align with providers' goals. For example, Chung et al. (2016) describe three types of SM – namely self-reflective, action-oriented and affective-oriented. It may be that while action-oriented SM may be more aligned with providers' expectations, self-reflective and affective-oriented SM can be very personal and align more with patient expectations. Lack of alignment between the patients' and providers' SM orientation may influence their interaction patterns.

Regarding the impacts of patient-provider co-management, some unexpected results are reported. ITSM may create new expectations on the patients' side in that they may expect more timely feedback from the clinicians and hope the clinicians show sympathy. When it takes time for the clinicians to formally assess the information, not knowing what SM will reveal creates negative feelings of uncertainty and anxiety in patients (Andersen et al. 2017). No studies examine the impact of patient-provider co-management on chronic care goal achievement.

1.6.4.3 Intermediate Outcome 3: Social Interactions with Family and Peers

Social interaction with family and peers has received relatively less attention compared to other intermediate outcomes, and yet the results are promising. Social functionality has become increasingly common for many commercialized ITSM devices such as fitness trackers, but in the medical context, it seems that sharing data and personal status with peers is seldom promoted by healthcare providers. Four studies report evidence regarding ITSM and social interaction (see Figure 1.9 and Table 1.14). Participants describe that the ability to share chronic care experiences and information through ITSM allows them to emotionally and instrumentally support – and be supported by – their peers (Fukuoka et al. 2011; Roblin 2011). One experiment that allowed both parents and adolescents to access ITSM found that parents' participation is positively associated with adolescents' SM rate and weight reduction (Tu et al. 2017). However, some patients feel reluctant to share with others as the patients may feel uncomfortable with how other people perceive them and do not want to attract attention to their disease (Kendall et al. 2015).

A number of studies purposefully incorporated social media or virtual communities as part of their complex ITSM interventions (e.g. Fukuoka et al. 2011; Glasgow et al. 2011; Jones et al. 2014; Partridge et al. 2016), providing IT affordances for sharing information with peers and families. However, formal investigation regarding the role of social interaction in a virtual environment, its delivery method, and its impacts, is particularly scant. Future research can further investigate this intermediate outcome as the preliminary results demonstrate the potential power of social and external support.

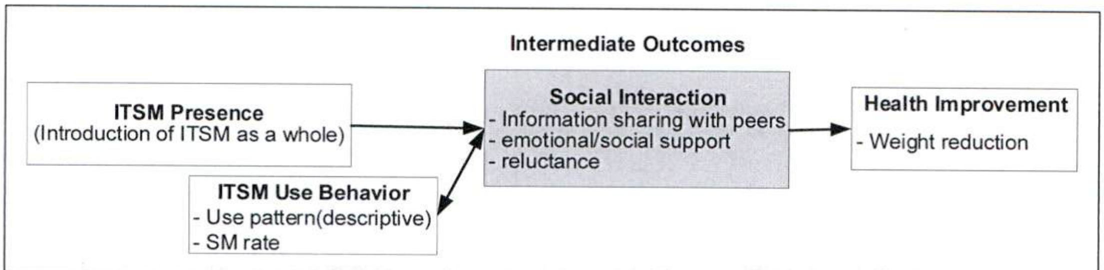


Figure 1.9 Relationships investigated for social interaction

Table 1.14 Role of social interaction

Impacts on Social Interactions with Peers from:	
ITSM presence	Fukuoka et al. (2011), Roblin (2011)
ITSM use	*[Kendall et al. (2015)]*
Effects of Social Interactions with Peers on:	
ITSM use	Tu et al. (2017)
Health outcome	Tu et al. (2017)

Note. Studies in italicized brackets have non-supportive or mixed results.

1.6.4.4 Intermediate outcome 4: Intervention Satisfaction and Compliance

Intervention compliance (or adherence) refers to the degree to which a patient correctly follows the intervention (also termed treatment, medication, or experiment depending on the study design) on schedule and as prescribed (Chakrabarti 2014). Intervention compliance may overlap with ITSM use (most often measured by SM frequency); however ITSM is often part of a larger complex intervention and involves specific instructions or targets beyond simple frequency. As the current practice of chronic care usually involves non-IT components for education, counseling and feedback, patients' compliance and satisfaction with the entire intervention can be important intermediate outcomes. Sixteen studies report relevant results (see Figure 1.10 and Table 1.15).

Participants generally report positive evaluations of the whole intervention (i.e. satisfaction and acceptability, Aguiar et al. 2017; Aharonovich et al. 2017; Goffinet et al. 2017; Pedersen et al. 2012; Rader et al. 2017; Steinberg et al. 2013). However, when assessing specific parts of the intervention, two studies report negative evaluations—Fukuoka et al. (2011) report negative perceptions due to strict SM data input timeframe, and Welch et al. (2013) report low perceived benefits regarding certain indicators (e.g., sodium and fluid adherence in diet SM). These findings imply that procedural barriers such as inflexible rules and insufficient understanding of the intervention purpose may impede intervention compliance and satisfaction. The actual compliance behavior, usually represented as attendance or participation, is influenced by intervention design. For example, intensive counseling may be more effective than light counseling in compelling participants to fully comply with the intervention (Allen et al. 2013); email or text prompts are a better approach for follow-up participation than phone-based reminders (Hall and Murchie 2014); and adding mobile SM facilitates patients' adherence to a given medical therapy as compared to standard care (Hostler et al. 2017).

Four studies examine health goal achievement as a result of intervention satisfaction and compliance. Patients' level of participation in the intervention program is associated with total weight loss and BMI reduction (Tu et al. 2017; Turner-McGrievy et al. 2017), but reduction in waist circumferences is not significant (Tu et al. 2017). Paula et al. (2017) found patients' perceptions of intervention benefits have a positive correlation with quality of life measures. Though the evidence is limited, it highlights the potential role of intervention satisfaction and compliance as an intermediating mechanism. Future research can further investigate the ITSM factors which influence intervention satisfaction and compliance as well as its impacts on chronic care goal achievement.

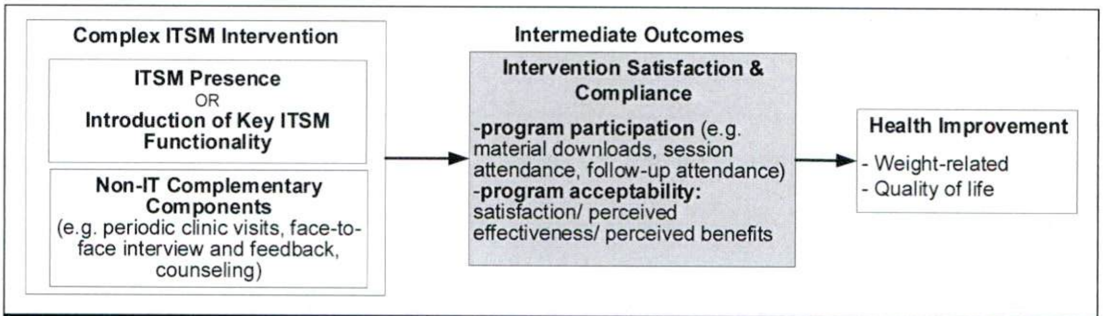


Figure 1.10 Relationships investigated for intervention acceptance

Table 1.15 Role of intervention satisfaction and compliance

Impacts on Intervention Satisfaction and Compliance from:	
Complex ITSM intervention	Allen et al. (2013), Pedersen et al. (2012), Steinberg et al. (2013), Abrantes et al. (2017), Aguiar et al. (2017), Aharonovich et al. (2017b), Hostler et al. (2017), Rader et al. (2017), Goffinet et al. (2017) *[Dennison et al. (2014), Fukuoka et al. (2011), Hall and Murchie (2014), Welch et al. (2013)]*
Effects of Intervention Satisfaction and Compliance on:	
Health improvement	Paula et al. (2017), Turner-McGrievy et al. (2017) *[Tu et al. (2017)]*

Note. Studies in italicized brackets have non-supportive or mixed results.

1.6.4.5 Theme 4 Discussion and Future Directions

Key takeaways and future research suggestions are presented in Table 1.16. By identifying and categorizing the key intermediate outcomes of ITSM, we find four intermediate outcomes that may help to facilitate ultimate chronic care goal achievement. As patients develop ITSM routines, their ability to learn from data improves, which may enhance their self-understanding and increase their beliefs in one's ability to perform health self-management actions, thus facilitating ultimate goal achievement. The social connection affordance provided by ITSM opens new opportunities for patients to share data with providers and peers through which they may obtain additional emotional and instrumental support, which is beneficial for successful behavior change and health improvement. From the providers' perspective, having access to patients' SM data gives them new sources of information, which is helpful in diagnosing and creating personalized treatment plans. Yet, IT functional barriers and patients' reluctance to share data may impede social interaction and patient-provider co-management. Since the intervention can be longitudinal and complex, overall satisfaction with the intervention may facilitate compliance, which is generally considered a necessary part of successful chronic care.

As in themes 2 and 3, the majority of the studies do not employ any theoretical lens to examine the relationships of theme 4. Despite the potential roles of these intermediate outcomes, existing studies only provide initial evidence, and most of them are qualitative and descriptive, without deep theoretical explanation. One notable exception is Chung and colleagues (2016) which draws on Lee's (2007) model of boundary negotiation artifacts to understand the importance of these artifacts in supporting patient-provider collaboration. In another study, a sociocultural perspective is used to highlight the potential for technology and media – of which ITSM is one example – to distort body image in patients with eating disorders (Eikey 2017). However, the general lack of theory in theme 4 means that many unknowns remain regarding how and why these intermediate outcomes arise. For example, we do not know if or how ITSM improves these intermediate outcomes such as self-efficacy and motivation, how self-understanding develops and influences chronic care goal achievement, or how patient-provider co-management is facilitated and constrained by ITSM functionalities. Future research related to theme 4 could draw on social representation theory (Wagner and Hayes 2005), the concept of IT identify (Carter 2015), or employ a practice lens (Feldman and Orlikowski 2011) to better understand the role of intermediate outcomes in ITSM (see below for a few illustrations).

In addition to the general lack of theory, these intermediate outcomes are proposed as mechanisms linking ITSM and goal achievement, but seldom do existing studies test the actual impacts. Thus, we do not know whether or not improving patient learning and interactions with providers and peers can indeed generate positive effects on goal achievement.

Accordingly, we propose three broad areas of future research. First, the role of intermediate outcomes is generally understudied, and more rigorous investigation is needed to investigate their mediating effects on chronic care goal achievement. For example, longitudinal research can be conducted in order to understand how ITSM influences patient-provider co-management, which in turn influences chronic care goal achievement. Patient SM practice can be relatively flexible, while clinical procedures can be very structured, creating reluctance on the part of providers if they do not trust the patients' collected data. Future research can focus on how ITSM promotes or constrains co-management, and a practice lens (e.g., Feldman and Orlikowski 2011) can be used to understand micro-processes that bring about the effects. Similarly, ITSM affordances that support social interaction with families and peers may change dramatically with the emergence of new communication and social media technologies. New technological support of virtual presence may influence how people behave and interact in virtual spaces with their sensitive health issues, which further influences chronic care goal achievement. Social representation theories (Wagner and Hayes 2005) and theories on IT identity (Carter 2015) can be used to explore how patients' virtual and illness-related identities influence information sharing and digital participation on ITSM-related social platforms.

Another interesting point is the mediating role of intervention compliance. Studies in other healthcare contexts have tested the mediating role of treatment compliance for various outcomes, some finding supportive evidence and others finding none (e.g. Ilgen et al. 2006; Turk et al. 2013; Wang et al. 2012). With the introduction of ITSM and the digitization of various treatment components previously delivered offline, the form of compliance may change, and future studies can investigate the ITSM factors, either technological or otherwise, that influence intervention satisfaction and compliance as well as their impacts on chronic care goal achievement.

Building upon this suggestion, future research can also investigate potential interactions of these intermediate outcomes. For example, patients' expectancy and perceived control may influence their further treatment adherence (Gonzalez et al. 2015; Westra et al. 2007). Investigation of how intermediate outcomes are connected and influence each other is also in line with the idea of affordance bundle and path dependence of affordance actualization in the theory (Strong et al. 2014). Since affordances are usually bundled, SM processes may have a cascading effect in that latter stages of ITSM intervention cannot be successfully performed if the former stages have not been accomplished (Li et al. 2010). Since users may deal with multiple affordances at the same time to achieve interrelated goals (e.g. successful self-assessment depends on accurate capturing of SM outcomes and meaningful data presentation), ITSM efficacy may largely depend on the emergence and actualization of these nested affordances. Thus, understanding how patients actualize the affordances (i.e. result in various intermediate outcomes) and their path dependence (i.e., the interaction between the actualized outcomes) has theoretical significance to unfold how ITSM promotes chronic care.

Finally, a feedback loop may exist during ITSM, which should be an important mechanism since it can further promote the actualization of additional intermediate outcomes. Although the affordance actualization framework suggests such feedback from actualized affordances (i.e. intermediate outcomes) to the affordance potentials, we did not find studies explicitly examining this type of relationship. However, several studies reported goal update in ITSM devices during the treatment, usually in a periodic manner (e.g. Abrantes et al. 2017; Painter et al. 2017). As per goal setting theories, goal specificity influences task execution strategy and performance, yet the level of ability, efficacy beliefs and outcome feedback also influence an individual's commitment to the goal and goal setting (Earley et al. 1990; Greenlees et al. 2000; Hollenbeck et al. 1989; Klein et al. 1999). Thus, feedback loops from intermediate outcomes (e.g. patient learning) and behavior change (e.g. physical activity performance) to ITSM use are possible. Future research can explore various possibilities of feedback mechanisms, including positive and negative reinforcement, and unfold the theoretical reasons behind them.

Table 1.16 Summary of Theme 4
What is known
• ITSM can help patients understand and learn from their data to make links between their daily activities and health conditions. • IT functionalities, such as automated data sharing, enhance patient-provider co-management beyond traditional periodic face-to-face data review. • Patient-provider co-management may be impeded by IT functional barriers and physician's mistrust of the system or data. • ITSM that incorporates social functionality can enable emotional and instrumental support from peers.

Table 1.16 Summary of Theme 4
• Intervention satisfaction and compliance is beneficial for achieving health goals (e.g. weight loss)
What is unknown and suggestions for future research
• Whether and how ITSM improves intermediate outcomes such as patient satisfaction, patient self-efficacy and motivation, patient awareness, patient-provider co-management, and social interaction with peers. • Whether these improvements in intermediate outcomes ultimately influence achievement of chronic care goals. ➢ Investigate and more rigorously test if these intermediate outcomes mediate the effects on ITSM use on chronic care goal achievement. ➢ Longitudinally examine how ITSM influences patient-provider co-management and the ultimate impacts on chronic care goal achievement. ➢ Further study the role of ITSM-enabled social interaction in achieving chronic care goals. ➢ Investigate the ITSM factors that influence intervention satisfaction and compliance, as well as their impacts on chronic care goal achievement. ➢ Investigate the potential interactions between these intermediate outcomes. • How to design ITSM interventions so that procedural barriers are minimized, and intervention satisfaction and compliance are improved. ➢ Investigate the impacts of new technologies on the actualization of intermediate outcomes. • Whether feedback loops actually exist and how they work. ➢ Investigate the multiple possible feedback mechanisms that could influence ITSM use.

1.7 Discussion

This paper reviews the literature on IT-based self-monitoring for chronic disease and develops a framework to help guide future research. Drawing on the affordance actualization framework (Strong et al., 2014), our synthesis focuses on four key themes: ITSM functionalities (that enable ITSM affordances), ITSM use and user experiences, intermediate outcomes, and chronic care goal achievement. The key findings find some support for the potential usefulness of ITSM – either presented as a standalone system or as part of a complex intervention – and its positive impacts on certain behavior change and health improvement outcomes, namely physical activity and weight reduction. Our synthesis also reveals three overarching issues related to research on ITSM for chronic care, which– along with related opportunities for future research – are outlined next (see Table 1.17).

Table 1.17 Overarching Research Issues and Future Research Suggestions	
Issue	Future Research Suggestions
Fragmentation of ITSM for chronic care research	Pursue a more complete approach connecting ITSM characteristics – through use and intermediate outcomes – to chronic care goal achievement. Use our framework to specify how future research adds to the ongoing investigation of ITSM for chronic care.
Shallow Understanding of the Role of IT	Pursue an in-depth understanding of the transformational role of IT in chronic care. For example: - Understand how ITSM transforms patient engagement in chronic care - Understand how patient-initiated ITSM transforms healthcare practices and the role of healthcare providers - Understand how ITSM transforms patient record management (e.g. integration of informal patient-generated information into standardized clinical information, issues of information quality)
Paucity of Strong Theory	Pursue more diverse perspectives of ITSM for chronic care. Pursue multi-level explanations of ITSM implementation, use and impacts. For example: - Cognitive and behavioral level explanations for human-IT interaction. - Interpersonal-level explanations for patient-provider co-management and peer-to-peer interaction. - System-level explanations for emerging attributes and capacities due to synergistic effects of various intermediate outcomes.

1.7.1 Research Issue 1- Fragmentation of ITSM for Chronic Care Research

ITSM for chronic care is multidisciplinary by nature in that people, IT and healthcare practices are intertwined pillars transforming chronic disease self-management. Drawing on these stakeholders, multiple intermediate outcomes emerged from the extant studies (patient learning and self-reflection, patient-provider co-management, social interactions with families and peers, and intervention satisfaction and compliance). These intermediate outcomes may serve as important mechanisms between ITSM use and chronic care goal achievement; however, additional examination is needed to reach definitive conclusions. Instead, the reviewed studies exhibit a fragmented landscape in which a large proportion of studies only examines the direct impacts of ITSM on goal achievement (ignoring the multiple mechanisms through which the ITSM impacts goal achievement) while another large proportion only examines ITSM design and mechanisms (ignoring ultimate impacts on goal achievement).

While individual studies may reasonably concentrate on a single aspect of this complex process (e.g. effects of ITSM design on use), too few studies take a more comprehensive approach that is necessary to build a solid chain of evidence connecting ITSM characteristics – through use and intermediate outcomes – to chronic care goal achievement. This limits the development of the field by restricting the definitive conclusions that can be drawn and the progress that can be made.

Our synthesized framework is a useful starting point for future research. It offers a more integrative understanding of ITSM for chronic care, and future research should take a broader focus by including concepts along the path from ITSM to ultimate impacts. It also helps future researchers identify areas of interest and specify how their research adds to the ongoing investigation of ITSM for chronic care.

1.7.2 Research Issue 2- Shallow Understanding of the Role of IT

The extant studies generally take a simplistic tool view of IT (for example, the presence or absence of ITSM). How IT can transform multiple aspects of chronic care has received little investigation. New ITSM technological developments – such as wearables, insideables, and complex AI – may have far-reaching impacts on patients and their healthcare practices. A shallow understanding of the role of IT may lead to missed opportunities, both in terms of practice and research. To illustrate, three emerging transformations entwined with technology developments are outlined for future examination.

Newer ITSM advances may transform patient engagement in chronic care. Current studies have examined ITSM use frequency and satisfaction, which are important parts of engagement. However, new insideable technologies may fundamentally shift how patients interact with ITSM devices. The meaning of *use frequency* is unclear when an implanted ITSM device is automatically transmitting data to the system and/or medical provider; patient engagement and use do not occur through action, but rather through inaction (e.g., by not turning off data capture functionality or by not removing an implanted device). Thus, patient engagement may not link to use frequency but may be more closely related to continued tolerance of the device. Rather than perspectives based on the theory of planned behavior, theories of decision inertia from behavioral economics may be the key to understanding patient engagement in future ITSM (Madrian and Shea 2001).

ITSM technological advancements also transform the role of the healthcare provider. Traditional paper-based SM was often initially recommended by the patient's healthcare provider based on specific needs related to the patient's chronic disease, and was accompanied with specific medical protocols. The explosion of access to ITSM devices by the general public (Gartner 2018) means that instead of generally being initiated by the provider, SM is now often initiated by the patients themselves. While there are many benefits to the wider diffusion of ITSM, the resulting SM practices may be less structured than provider-initiated SM, may deviate from disease-specific SM protocols, and at times may lead to practices that are sub-optimal or not recommended for the patient's chronic disease (Gabriels and Moerenhout 2018). Thus, provider influence and control over the SM process may be diminished as compared to paper-based SM or ITSM with earlier technologies.

The transformational effects of ITSM on patient-generated information also create opportunities and challenges for data management and use. Whereas traditional clinical information is often standardized, structured, formal, and gathered according to specific protocols (e.g., measurements taken at regular intervals using verified measurement tools, de Vet 2003), patients' SM data are often unstandardized, unstructured, less formal, and gathered in an ad-hoc way. While patient-generated SM information can complement the more traditional patient health records, there are many obstacles to integrating the two. The clinical infrastructure and practices may not support the storage and analysis of patients' SM data, thus healthcare professionals may consider it as extra work, or they may not have the relevant skills to proficiently analyze these data and incorporate them into personalized treatment plans. Moreover, information quality can be a major issue as patient-generated information may not be reliable enough to support formal clinical processes (West et al. 2017). Finally, the significant amount of data created by ITSM may make it difficult for physicians – who often have hundreds of patients to follow – to closely monitor the SM data and appropriately adjust their clinical recommendations. Emerging AI techniques have the potential to alleviate the pressure on busy physicans by helping them monitor patients' data (e.g. auto-detection of anomalies and unexpected deviations) and make treatment decisions (e.g., automatic diagnoses and proactive interventions based on data trends). Thus, future research should investigate patient health record management issues caused by the technological advances and the integration of patient SM data.

1.7.3 Research Issue 3- Paucity of Strong Theory

Our examination of the literature shows that, in all parts of our overarching framework, ITSM for chronic care research is not theory-rich. One hundred studies (63% of the sample) do not use any theory or develop new theories, and those studies that do use theory do not contribute back to or extend the original theory. The majority of these studies cite theory to inform interventions, tool design, or measure development, but do not use theory to explain the relationships under investigation. The main theories used and their corresponding studies are listed in Table 1.18.

Many of the medical studies focus on description and prediction rather than on explanation: this may be appropriate since the purpose of much healthcare research is evaluating intervention effectiveness rather than contributing to theory. However, abstracting to a theoretical understanding is also important. Given the fast pace of technological change and the complexity of the healthcare ecosystem, a theoretical understanding of ITSM's abstracted functionalities and the underlying causal mechanisms for their effects enables the accumulation of knowledge about ITSM for chronic disease management and avoids a plethora of piecemeal and fragmented studies.

Table 1.18 Theory Used in the Extant Studies	
Theory*	**Studies that reference the theory**
Social cognitive theory	Allen et al. (2013), Bonilla et al. (2015), Cadmus-Bertram et al. (2015), Dorsch et al. (2015), Fukuoka et al. (2011), Kendall et al. (2015), Schroder (2011), Stark et al. (2011), Turk et al. (2013), Abrantes et al. (2017), Hales et al. (2017), Jakicic et al. (2016), Mummah et al. (2017), Plow and Golding (2017), Tu et al. (2017)
Behavior change theories	Ryan et al. (2012), Stark et al. (2011), Aguiar et al. (2017), Cai et al. (2017), Chen et al. (2017), Hostler et al. (2017), Mummah et al. (2017), Plow and Golding (2017), Tu et al. (2017)
Control theory	Spring et al. (2017), Kendall et al. (2015), Kolodziejczyk et al. (2014), Schroder (2011)
Theory of planned behavior and extended theories	Laing et al. (2015), Stark et al. (2011), Stark et al. (2011), Storni (2010), Biddle et al. (2017)
Self-efficacy theory	Fukuoka et al. (2011), Laing et al. (2015), Rader et al. (2017), Izawa et al. (2006)
Chronic care model**	Tu et al. 2017, Karhula et al. 2015, Partridge et al. 2016, Roblin 2011
Cognitive behavioral therapy**	Barakat et al. 2017, Mantani et al. 2017, Zhu et al. 2017, Acharya et al. 2011, Naylor et al. 2008, Nicklas et al. 2014, Zhu et al. 2017

* Only theories that were used in more than one study are included in this table.

**The chronic care model and cognitive behavioral therapy are not technically theories, but treatment frameworks that were used in the extant studies to inform intervention design.

Based on our framework and literature synthesis, several potential avenues for using additional theoretical lenses to deepen our understanding of ITSM for chronic care are proposed. The future research examples provided are by no means exhaustive as our purpose is not to outline all of the relevant opportunities, but to illustrate a few key potential avenues.

As demonstrated in our review, few ITSM studies take a comprehensive approach that is necessary to understanding how ITSM characteristics – through use and intermediate outcomes – influence chronic care goal achievement. Thus, little is known about how and why ITSM effects occur. Moreover, the majority of studies take a deterministic view of the interventions without more micro- and in-depth examinations of users, user behaviors, or interpersonal interactions. While social cognitive theories and cognitive behavior therapy are often referenced in these studies, this narrow focus can only investigate a limited range of phenomena and research questions. Thus, more diverse perspectives of ITSM for chronic care are needed.

One starting point could be diversifying the level of explanation. First, instead of focusing on ITSM interventions, research could take a more micro-level approach. In theme 2, some patterns of temporary spikes and declines in ITSM use emerged. Future research could use theories related to goals and motivation (e.g. Locke 1991; Ryan and Deci 2000) to perform micro-level longitudinal investigations of the ebb and flow of ITSM use and how it is related to chronic care goal attainment (e.g., when chronic care goals are reached, does ITSM use continue, stop, or continue sporadically?). Research could also draw on theories of attitude change and affect to explore the interactions between patients' cognitions, affect and use behaviors (e.g. Anderson 1971; Maddux and Rogers 1983; Zhang 2013). For example, various unintended negative consequences of ITSM use emerged in theme 2, such as feeling overwhelmed, finding sub-optimal workarounds, and overuse. Research could draw on coping theory (e.g. Bhattacherjee et al. 2018; Stein et al. 2015) to understand how and why these unintended negative consequences emerge and how patients manage them. For example, the repeated visualization of one's own data related to chronic disease may act as a constant negative reminder, creating a type of stress and causing patients to ruminate too much on their health issues and perform unintended impulsive behaviors which may, in turn, lead to negative outcomes.

A second way of diversifying the level of explanation is to focus more on interpersonal interactions. ITSM often involves families, peers and healthcare providers, and the interactions between these groups are only rarely examined in the literature (see Table 1.13 and 1.14 for exceptions). For example, patient-provider interactions are potentially a key outcome driving chronic care goal achievement and should receive more research attention. A Foucauldian perspective (Foucault 1980; 1982) – which simultaneously considers knowledge, power, and practices as well as interactions between all three – could be used to provide a deeper understanding of how ITSM can change the power dynamics between patients and providers. During ITSM, patients produce various types of self-related knowledge. This knowledge is produced (but also constrained) by ITSM practices. The acquisition of this self-knowledge may change the power dynamics between patients and health providers. Alternatively, providers may use knowledge gleaned from shared ITSM data as a way to exert influence over patients. The effect of ITSM on the power dynamics of the patient-provider relationship has received little research attention.

Finally, we can take a system-level perspective, conceptualizing ITSM as a system within which various regulative mechanisms exist in different stages of SM in order to organize the interacting entities (e.g. patients and providers) and activities (e.g. SM data capture and reflection). By taking a systems perspective (e.g. Bailey 1994; Bertalanffy 1973), we allow the emergence of new properties (e.g. IT identity) that are the result of the synergistic effects of the structures (e.g. ITSM functionalities) and intermediate processes (e.g. patient learning and data sharing). For example, researchers could borrow the key concepts and principles from control theories (e.g. Hirschi 2017) to investigate the dimensions and specifications of the potential control mechanisms (e.g. ITSM rules imposed by IT infrastructure and physician instructions) and the impacts on ITSM outcomes. The advantage of a system perspective is that it can potentially provide an integrative understanding of ITSM as a whole system, yet is dynamic enough to allow the emergence of new mechanisms and attributes.

1.7.4 Limitations and Conclusions

This review of the ITSM literature outlined opportunities for future research in which more diverse perspectives can contribute to our understanding of the phenomenon. As noted above, the suggestions provided are by no means exhaustive. Our purpose was to not only recommend specific research questions and theoretical lenses, but to highlight overall directions for future research in order to diversify the phenomena under investigation. Nevertheless, there are several limitations related to our systematic review.

First, our review included literature published between 2006 and 2017. While studies published before 2006 examine self-monitoring for chronic disease management, the few that examine ITSM involve capabilities that are not comparable with recent IT advances. Second, although seeking to include as many relevant studies as possible, we only incorporate studies that explicitly mentioned our search terms in the title or abstract, which may have limited our sample pool. Some healthcare studies use terms such as "web-based intervention" or "mobile-based intervention" in the title or abstract without mentioning our search terms. However, to keep the number of screened articles to a manageable size (we screened the titles and abstracts of 5,152 articles), we did not expand our search to cover these more general terms. Third, we only included studies that explicitly incorporated chronic disease in their research objectives. We excluded studies investigating healthy behaviors for general populations (e.g. SM of physical activity for the general population without any explicit research objectives related to chronic disease management). While these studies can be related to health promotion, and the implication may be applicable to the chronic care context, they do not fall under our definition of chronic disease self-management. Future researchers may want to draw on this related work as there is some overlap in the types of data being monitored and technologies being used.

In conclusion, our synthesis shows that ITSM has the potential to help people manage their chronic diseases. However, additional studies are needed to address the research gaps outlined for each of the themes above and to address the three overarching issues in this field of research.

References

*Abrantes, A. M., Blevins, C. E., Battle, C. L., Read, J. P., Gordon, A. L., and Stein, M. D. 2017. "Developing a Fitbit-Supported Lifestyle Physical Activity Intervention for Depressed Alcohol Dependent Women," *Journal of Substance Abuse Treatment* (80), pp. 88-97 (doi: 10.1016/j.jsat.2017.07.006).

*Acharya, S. D., Elci, O. U., Sereika, S. M., Styn, M. A., and Burke, L. E. 2011. "Using a Personal Digital Assistant for Self-Monitoring Influences Diet Quality in Comparison to a Standard Paper Record among Overweight/Obese Adults," *Journal of the American Dietetic Association* (111:4), pp. 583-588 (doi: 10.1016/j.jada.2011.01.009).

*Adams, P., Murnane, E. L., Elfenbein, M., Wethington, E., and Gay, G. 2017. "Supporting the Self-Management of Chronic Pain Conditions with Tailored Momentary Self-Assessments," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver, Colorado, USA: ACM, pp. 1065-1077 (doi:10.1145/3025453.3025832).

Aggarwal, R., Kryscynski, D., Midha, V., and Singh, H. 2015. "Early to Adopt and Early to Discontinue: The Impact of Self-Perceived and Actual It Knowledge on Technology Use Behaviors of End Users," *Information Systems Research* (26:1), pp. 127-144 (doi: 10.1287/isre.2014.0564).

*Aguiar, E. J., Morgan, P. J., Collins, C. E., Plotnikoff, R. C., Young, M. D., and Callister, R. 2017. "Process Evaluation of the Type 2 Diabetes Mellitus Pulse Program Randomized Controlled Trial: Recruitment, Engagement, and Overall Satisfaction," *American Journal of Men's Health* (11:4), pp. 1055-1068 (doi: 10.1177/1557988317701783).

*Aharonovich, E., Hatzenbuehler, M., Johnston, B., O'Leary, A., Morgenstern, J., Wainberg, M., Yao, P., Helzer, J., and Hasin, D. 2006. "A Low-Cost, Sustainable Intervention for Drinking Reduction in the HIV Primary Care Setting," *AIDS Care* (18:6), pp. 561-568 (doi: 10.1080/09540120500264134).

*Aharonovich, E., Sarvet, A., Stohl, M., DesJarlais, D., Tross, S., Hurst, T., Urbina, A., and Hasin, D. 2017. "Reducing Non-Injection Drug Use in Hiv Primary Care: A Randomized Trial of Brief Motivational Interviewing, with and without Healthcall, a Technology-Based Enhancement," *Journal of Substance Abuse Treatment* (74), pp. 71-79 (doi: 10.1016/j.jsat.2016.12.009).

*Aharonovich, E., Stohl, M., Cannizzaro, D., and Hasin, D. 2017. "Healthcall Delivered via Smartphone to Reduce Co-occurring Drug and Alcohol Use in HIV-Infected Adults: A Randomized Pilot Trial," *Journal of Substance Abuse Treatment* (83), pp. 15-26 (doi: 10.1016/j.jsat.2017.09.013).

Alderson, P. 1998. "Theories in Health Care and Research: The Importance of Theories in Health Care," *BMJ: British Medical Journal* (317:7164), pp. 1007-1010.

*Allen, J. K., Stephens, J., Dennison Himmelfarb, C. R., Stewart, K. J., and Hauck, S. 2013. "Randomized Controlled Pilot Study Testing Use of Smartphone Technology for Obesity Treatment," *Journal of Obesity* (2013), Article ID 151597, pp. 1-7 (doi: 10.1155/2013/151597).

Altman, I. 1977. "Privacy Regulation: Culturally Universal or Culturally Specific?," *Journal of Social Issues* (33:3), pp. 66-84 (https://doi.org/10.1111/j.1540-4560.1977.tb01883.x).

*Ambeba, E. J., Ye, L., Sereika, S. M., Styn, M. A., Acharya, S. D., Sevick, M. A., Ewing, L. J., Conroy, M. B., Glanz, K., Zheng, Y., Goode, R. W., Mattos, M., and Burke, L. E. 2015. "The Use of Mhealth to Deliver Tailored Messages Reduces Reported Energy and Fat Intake," *Journal of Cardiovascular Nursing* (30:1), pp. 35-43 (doi: 10.1097/JCN.0000000000000120).

Anderson, N. H. 1971. "Integration Theory and Attitude Change," *Psychological Review* (78:3), pp. 171-206 (https://psycnet.apa.org/doi/10.1037/h0030834).

*Andersen, T. O., Andersen, P. R. D., Kornum, A. C., and Larsen, T. M. 2017. "Understanding Patient Experience: A Deployment Study in Cardiac Remote Monitoring," in *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*. Barcelona, Spain: ACM, pp. 221-230 (doi:10.1145/3154862.3154868).

Anderson, C., and Robey, D. 2017. "Affordance Potency: Explaining the Actualization of Technology Affordances," *Information and Organization* (27:2), pp. 100-115 (doi: 10.1016/j.infoandorg.2017.03.002).

*Ayobi, A., Marshall, P., Cox, A. L., and Chen, Y. 2017. "Quantifying the Body and Caring for the Mind: Self-Tracking in Multiple Sclerosis," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver, Colorado, USA: ACM, pp. 6889-6901 (doi:10.1145/3025453.3025869).

Bailey, K. D. 1994. *Sociology and the New Systems Theory: Toward a Theoretical Synthesis*, Albany, NY: SUNY Press.

Bandura, A. 1989. "Human Agency in Social Cognitive Theory," *American Psychologist* (44:9), pp. 1175-1184.

Bandura, A. 1991. "Social Cognitive Theory of Self-Regulation," *Organizational Behavior and Human Decision Processes* (50:2), pp. 248-287 (doi: 10.1016/0749-5978(91)90022-L).

Bandura, A. 1998. "Health Promotion from the Perspective of Social Cognitive Theory," *Psychology and Health* (13:4), pp. 623-649 (https://doi.org/10.1080/08870449808407422).

*Barakat, S., Maguire, S., Surgenor, L., Donnelly, B., Miceska, B., Fromholtz, K., Russell, J., Hay, P., and Touyz, S. 2017. "The Role of Regular Eating and Self-Monitoring in the Treatment of Bulimia Nervosa: A Pilot Study of an Online

Guided Self-Help CBT Program," *Behavioral Sciences* (7:3), p. 39 (doi: 10.3390/bs7030039.

Barr, V. J., Robinson, S., Marin-Link, B., Underhill, L., Dotts, A., Ravensdale, D., and Salivaras, S. 2003. "The Expanded Chronic Care Model: An Integration of Concepts and Strategies from Population Health Promotion and the Chronic Care Model," *Healthcare Quarterly* (7:1), pp. 73-82 (doi:10.12927/hcq.2003.16763).

Bartholomew, L. K., Parcel, G. S., Swank, P. R., and Czyzewski, D. I. 1993. "Measuring Self-Efficacy Expectations for the Self-Management of Cystic Fibrosis," *Chest* (103:5), pp. 1524-1530 (https://doi.org/10.1378/chest.103.5.1524).

*Bauer, M., Glenn, T., Grof, P., Rasgon, N., Alda, M., Marsh, W., Sagduyu, K., Schmid, R., Adli, M., and Whybrow, P. C. 2009. "Comparison of Sleep/Wake Parameters for Self-Monitoring Bipolar Disorder," *Journal of Affective Disorders* (116:3), pp. 170-175 (doi: 10.1016/j.jad.2008.11.014).

Bauer, U. E., Briss, P. A., Goodman, R. A., and Bowman, B. A. 2014. "Prevention of Chronic Disease in the 21st Century: Elimination of the Leading Preventable Causes of Premature Death and Disability in the USA," *The Lancet* (384:9937), pp. 45-52 (https://doi.org/10.1016/S0140-6736(14)60648-6).

Beaudry, A., and Pinsonneault, A. 2005. "Understanding User Responses to Information Technology a Coping Model of User Adaptation," *MIS Quarterly* (29:3), pp. 493-524.

*Berry, D. L., Blonquist, T. M., Patel, R. A., Halpenny, B., and McReynolds, J. 2015. "Exposure to a Patient-Centered, Web-Based Intervention for Managing Cancer Symptom and Quality of Life Issues: Impact on Symptom Distress," *Journal of Medical Internet Research* (17:6), p. e136 (doi: 10.2196/jmir.4190).

Bertalanffy, L. v. 1973. *General System Theory: Foundations, Development, Applications*, New York, NY: G. Braziller.

Bhattacherjee, A., Davis, C. J., Connolly, A. J., and Hikmet, N. 2018. "User Response to Mandatory IT Use: A Coping Theory Perspective," *European Journal of Information Systems* (27:4), pp. 395-414 (doi: 10.1057/s41303-017-0047-0).

*Biddle, S. J. H., Edwardson, C. L., Gorely, T., Wilmot, E. G., Yates, T., Nimmo, M. A., Khunti, K., and Davies, M. J. 2017. "Reducing Sedentary Time in Adults at Risk of Type 2 Diabetes: Process Evaluation of the Stand (Sedentary Time and Diabetes) RCT," *BMC Public Health* (17:1), p. 80 (doi: 10.1186/s12889-016-3941-9).

Biener, L., and Abrams, D. B. 1991. "The Contemplation Ladder: Validation of a Measure of Readiness to Consider Smoking Cessation," *Health Psychology* (10:5), pp. 360-365 (https://psycnet.apa.org/doi/10.1037/0278-6133.10.5.360).

Bodenheimer, T., Lorig, K., Holman, H., and Grumbach, K. 2002. "Patient Self-Management of Chronic Disease in Primary Care," *JAMA* (288:19), pp. 2469-2475 (doi:10.1001/jama.288.19.2469).

Bowlby, J. 1983. "Attachment and Loss," *Attachment* (1), Basic Books Classics.

*Bonilla, C., Brauer, P., Royall, D., Keller, H., Hanning, R. M., and DiCenso, A. 2015. "Use of Electronic Dietary Assessment Tools in Primary Care: An Interdisciplinary Perspective," *BMC Medical Informatics & Decision Making* (15:14). (doi: 10.1186/s12911-015-0138-6).

*Boyd, L. E., Jiang, X. L., and Hayes, G. R. 2017. "ProCom: Designing and Evaluating a Mobile and Wearable System to Support Proximity Awareness for People with Autism," in: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver, Colorado, USA: ACM, pp. 2865-2877 (doi:10.1145/3025453.3026014).

Brazier, J. E., Harper, R., Jones, N., O'cathain, A., Thomas, K., Usherwood, T., and Westlake, L. 1992. "Validating the SF-36 Health Survey Questionnaire: New Outcome Measure for Primary Care," *BMJ: British Medical Journal* (305:6846), pp. 160-164 (https://doi.org/10.1136/bmj.305.6846.160).

*Burke, L. E., Styn, M. A., Sereika, S. M., Conroy, M. B., Ye, L., Glanz, K., Sevick, M. A., and Ewing, L. J. 2012. "Using mHealth Technology to Enhance Self-Monitoring for Weight Loss a Randomized Trial," *American Journal of Preventive Medicine* (43:1), pp. 20-26 (doi: 10.1016/j.amepre.2012.03.016).

*Caballero-Ruiz, E., Garcia-Saez, G., Rigla, M., Villaplana, M., Pons, B., and Hernando, M. E. 2017. "A Web-Based Clinical Decision Support System for Gestational Diabetes: Automatic Diet Prescription and Detection of Insulin Needs," *International Journal of Medical Informatics* (102), pp. 35-49 (doi:10.1016/j.ijmedinf.2017.02.014).

Cade, J. E. 2017. "Measuring Diet in the 21st Century: Use of New Technologies," *Proceedings of the Nutrition Society* (76:3), pp. 276-282 (doi: 10.1017/S0029665116002883).

*Cadmus-Bertram, L., Marcus, B. H., Patterson, R. E., Parker, B. A., and Morey, B. L. 2015. "Use of the Fitbit to Measure Adherence to a Physical Activity Intervention among Overweight or Obese, Postmenopausal Women: Self-Monitoring Trajectory during 16 Weeks," *JMIR mHealth and uHealth* (3:4), p. e96 (doi: 10.2196/mhealth.4229).

*Cadmus-Bertram, L., Wang, J. B., Patterson, R. E., Newman, V. A., Parker, B. A., and Pierce, J. P. 2013. "Web-Based Self-Monitoring for Weight Loss among Overweight/Obese Women at Increased Risk for Breast Cancer: The HELP Pilot Study," *Psycho-Oncology* (22:8), pp. 1821-1828 (doi: 10.1002/pon.3219).

*Cai, R. A., Beste, D., Chaplin, H., Varakliotis, S., Suffield, L., Josephs, F., Sen, D., Wedderburn, L. R., Ioannou, Y., Hailes, S., and Eleftheriou, D. 2017. "Developing and Evaluating JIApp: Acceptability and Usability of a Smartphone

App System to Improve Self-Management in Young People with Juvenile Idiopathic Arthritis," *JMIR mHealth and uHealth* (5:8), p. e121 (doi: 10.2196/mhealth.7229).

*Carels, R. A., Selensky, J. C., Rossi, J., Solar, C., and Hlavka, R. 2017. "A Novel Stepped-Care Approach to Weight Loss: The Role of Self-Monitoring and Health Literacy in Treatment Outcomes," *Eating Behaviors* (26), pp. 76-82 (doi: 10.1016/j.eatbeh.2017.01.009).

Carter, M. 2015. "Me, my self, and I (T): conceptualizing information technology identity and its implications, " *MIS Quarterly* (39:4), pp. 931-957.

*Carter, M. C., Burley, V. J., Nykjaer, C., and Cade, J. E. 2013. "Adherence to a Smartphone Application for Weight Loss Compared to Website and Paper Diary: Pilot Randomized Controlled Trial," *Journal of Medical Internet Research* (15:4), p. e32 (doi: 10.2196/jmir.2283).

CCS Insight. 2017. "*Wearables Market to Be Worth $25 Billion by 2019*," Release from http://www.ccsinsight.com/press/company-news/2332-wearables-market-to-be-worth-25-billion-by-2019-reveals-ccs-insight.

de Vet, H. C., Terwee, C. B., and Bouter, L. M. 2003. "Current Challenges in Clinimetrics," *Journal of clinical epidemiology* (56:12), pp. 1137-1141.

CDC. 2016. "*Chronic Disease Overview*," Release from https://www.cdc.gov/chronicdisease/overview/index.htm.

Chakrabarti, S. 2014. "What's in a Name? Compliance, Adherence and Concordance in Chronic Psychiatric Disorders," *World Journal of Psychiatry* (4:2), pp. 30-36 (doi: 10.5498/wjp.v4.i2.30).

*Chambliss, H. O., Huber, R. C., Finley, C. E., McDoniel, S. O., Kitzman-Ulrich, H., and Wilkinson, W. J. 2011. "Computerized Self-Monitoring and Technology-Assisted Feedback for Weight Loss with and without an Enhanced Behavioral Component," *Patient Education and Counseling* (85:3), pp. 375-382 (https://doi.org/10.1016/j.pec.2010.12.024).

Chen, C.-M., and Yeh, M. C. 2015. "The Experiences of Diabetics on Self-Monitoring of Blood Glucose: A Qualitative Metasynthesis," *Journal of Clinical Nursing* (24:5-6), pp. 614-626 (https://doi.org/10.1111/jocn.12691).

*Chen, J., Lieffers, J., Bauman, A., Hanning, R., and Allman-Farinelli, M. 2017. "The Use of Smartphone Health Apps and Other Mobile Health (mHealth) Technologies in Dietetic Practice: A Three Country Study," *Journal of Human Nutrition and Dietetics* (30:4), pp. 439-452 (doi: 10.1111/jhn.12446).

Chomutare, T., Fernandez-Luque, L., Arsand, E., and Hartvigsen, G. 2011. "Features of Mobile Diabetes Applications: Review of the Literature and Analysis of Current Applications Compared against Evidence-Based Guidelines," *Journal of Medical Internet Research* (13:3), p. e65 (doi: 10.2196/jmir.1874).

*Chung, C.-F., Cook, J., Bales, E., Zia, J., and Munson, S. A. 2015. "More Than Telemonitoring: Health Provider Use and Nonuse of Life-Log Data in Irritable Bowel Syndrome and Weight Management," *Journal of Medical Internet Research* (17:8), p. e203 (doi: 10.2196/jmir.4364).

*Chung, C.-F., Dew, K., Cole, A., Zia, J., Fogarty, J., Kientz, J. A., and Munson, S. A. 2016. "Boundary Negotiating Artifacts in Personal Informatics: Patient-Provider Collaboration with Patient-Generated Data," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. San Francisco, California, USA: ACM, pp. 770-786 (doi:10.1145/2818048.2819926).

Coleman, K., Austin, B. T., Brach, C., and Wagner, E. H. 2009. "Evidence on the Chronic Care Model in the New Millennium," *Health Affairs* (28:1), pp. 75-85 (doi: 10.1377/hlthaff.28.1.75).

*Conroy, M. B., Yang, K., Elci, O. U., Gabriel, K. P., Styn, M. A., Wang, J., Kriska, A. M., Sereika, S. M., and Burke, L. E. 2011. "Physical Activity Self-Monitoring and Weight Loss: 6-Month Results of the Smart Trial," *Medicine and Science in Sports and Exercise* (43:8), pp. 1568-1574 (doi: 10.1249/MSS.0b013e31820b9395).

*Coppini, G., Zuccala, V. C., De Maria, R., Nazare, J. A., Morales, M. A., and Colantonio, S. 2017. "User Acceptance of Self-Monitoring Technology to Prevent Cardio-Metabolic Diseases: The Wize Mirror," in *Proceedings of the 2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. Rome, Italy: IEEE, pp. 265-271 (doi: 10.1109/WiMOB.2017.8115837).

*Cosson, E., Baz, B., Gary, F., Pharisien, I., Nguyen, M. T., Sandre-Banon, D., Jaber, Y., Cussac-Pillegand, C., Banu, I., Carbillon, L., and Valensi, P. 2017. "Poor Reliability and Poor Adherence to Self-Monitoring of Blood Glucose Are Common in Women with Gestational Diabetes Mellitus and May Be Associated with Poor Pregnancy Outcomes," *Diabetes Care* (40:9), pp. 1181-1186 (doi: 10.2337/dc17-0369).

*Cushing, C. C., Jensen, C. D., and Steele, R. G. 2011. "An Evaluation of a Personal Electronic Device to Enhance Self-Monitoring Adherence in a Pediatric Weight Management Program Using a Multiple Baseline Design," *Journal of Pediatric Psychology* (36:3), pp. 301-307 (doi: 10.1093/jpepsy/jsq074).

Cuttone, A., Lehmann, S., and Larsen, J. E. 2013. "A Mobile Personal Informatics System with Interactive Visualizations of Mobility and Social Interactions," in *Proceedings of the 1st ACM International Workshop on Personal Data Meets Distributed Multimedia*. Barcelona, Spain: ACM, pp. 27-30 (doi: 10.1145/2509352.2509397).

Daley, A. J., and Duda, J. L. 2006. "Self-Determination, Stage of Readiness to Change for Exercise, and Frequency of Physical Activity in Young People," *European*

Journal of Sport Science (6:4), pp. 231-243 (https://doi.org/10.1080/17461390601012637).

Deci, E. L., and Ryan, R. M. 2008. "Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health," *Canadian Psychology* (49:3), pp. 182-185 (https://psycnet.apa.org/doi/10.1037/a0012801).

*Dennison, L., Morrison, L., Lloyd, S., Phillips, D., Stuart, B., Williams, S., Bradbury, K., Roderick, P., Murray, E., Michie, S., Little, P., and Yardley, L. 2014. "Does Brief Telephone Support Improve Engagement with a Web-Based Weight Management Intervention? Randomized Controlled Trial," *Journal of Medical Internet Research* (16:3), p. e95 (doi: 10.2196/jmir.3199).

Deterding, S. 2015. "The Lens of Intrinsic Skill Atoms: A Method for Gameful Design," *Human–Computer Interaction* (30:3-4), pp. 294-335 (https://doi.org/10.1080/07370024.2014.993471).

*Di Bartolo, P., Nicolucci, A., Cherubini, V., Iafusco, D., Scardapane, M., and Rossi, M. 2017. "Young Patients with Type 1 Diabetes Poorly Controlled and Poorly Compliant with Self-Monitoring of Blood Glucose: Can Technology Help? Results of the i-NewTrend Randomized Clinical Trial," *Acta Diabetologica* (54:4), pp. 393-402 (doi: 10.1007/s00592-017-0963-4).

DiClemente, C. C., Schlundt, D., and Gemmell, L. 2004. "Readiness and Stages of Change in Addiction Treatment," *The American Journal on Addictions* (13:2), pp. 103-119 (doi: 10.1080/10550490490435777).

*Dietrich, J. E., Yee, D. L., Santos, X. M., Bercaw-Pratt, J. L., Kurkowski, J., Soni, H., Lee-Kim, Y. J., Shah, M. D., Mahoney, D., and Srivaths, L. V. 2017. "Assessment of an Electronic Intervention in Young Women with Heavy Menstrual Bleeding," *Journal of Pediatric and Adolescent Gynecology* (30:2), pp. 243-246. (doi: 10.1016/j.jpag.2016.10.006).

*Donaldson, J. M., and Normand, M. P. 2009. "Using Goal Setting, Self-Monitoring, and Feedback to Increase Calorie Expenditure in Obese Adults," *Behavioral Interventions* (24:2), pp. 73-83 (https://doi.org/10.1002/bin.277).

*Dorsch, M. P., Farris, K. B., Bleske, B. E., and Koelling, T. M. 2015. "A Web Application for Self-Monitoring Improves Symptoms in Chronic Systolic Heart Failure," *Telemedicine and e-Health* (21:4), pp. 267-271 (doi: 10.1089/tmj.2014.0095).

*Dowell, S. A., and Welch, J. L. 2006. "Use of Electronic Self-Monitoring for Food and Fluid Intake: A Pilot Study," *Nephrology Nursing Journal* (33:3), pp. 271-277.

*Downing, J., Bollyky, J., and Schneider, J. 2017. "Use of a Connected Glucose Meter and Certified Diabetes Educator Coaching to Decrease the Likelihood of Abnormal Blood Glucose Excursions: The Livongo for Diabetes Program," *Journal of Medical Internet Research* (19:7), p. e234 (doi: 10.2196/jmir.6659).

Earley, P. C., Northcraft, G. B., Lee, C., and Lituchy, T. R. 1990. "Impact of Process and Outcome Feedback on the Relation of Goal Setting to Task Performance," *The Academy of Management Journal* (33:1), pp. 87-105 (doi: 10.2307/256353).

Echterhoff, G., Higgins, E. T., and Levine, J. M. 2009. "Shared Reality: Experiencing Commonality with Others' Inner States about the World," *Perspectives on Psychological Science* (4:5), pp. 496-521 (doi: 10.1111/j.1745-6924.2009.01161.x).

*Edge, J., Acerini, C., Campbell, F., Hamilton-Shield, J., Moudiotis, C., Rahman, S., Randell, T., Smith, A., and Trevelyan, N. 2017. "An Alternative Sensor-Based Method for Glucose Monitoring in Children and Young People with Diabetes," *Archives of Disease in Childhood* (102:6), pp. 543-549 (doi: 10.1136/archdischild-2016-311530).

*Eikey, E. V., and Reddy, M. C. 2017. "It's Definitely Been a Journey": A Qualitative Study on How Women with Eating Disorders Use Weight Loss Apps," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver, Colorado, USA: ACM, pp. 642-654 (doi: 10.1145/3025453.3025591).

Epstein, D. A. 2015. "Personal Informatics in Everyday Life," in *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. Osaka, Japan: ACM, pp. 429-434 (doi: 10.1145/2800835.2801643).

Epstein, R. M., Siegel, D. J., and Silberman, J. 2008. "Self-Monitoring in Clinical Practice: A Challenge for Medical Educators," *Journal of Continuing Education in the Health Professions* (28:1), pp. 5-13 (doi: 10.1002/chp.149).

Fairburn, C. G., and Rothwell, E. R. 2015. "Apps and Eating Disorders: A Systematic Clinical Appraisal," *The International Journal of Eating Disorders* (48:7), pp. 1038-1046 (doi: 10.1002/eat.22398).

Farmer, A., Wade, A., Goyder, E., Yudkin, P., French, D., Craven, A., Holman, R., Kinmonth, A. L., and Neil, A. 2007. "Impact of Self-Monitoring of Blood Glucose in the Management of Patients with Non-Insulin Treated Diabetes: Open Parallel Group Randomised Trial," *BMJ: British Medical Journal* (335:7611), pp. 132-136 (doi: 10.1136/bmj.39247.447431.BE).

*Faurholt-Jepsen, M., Frost, M., Ritz, C., Christensen, E. M., Jacoby, A. S., Mikkelsen, R. L., Knorr, U., Bardram, J. E., Vinberg, M., and Kessing, L. V. 2015. "Daily Electronic Self-Monitoring in Bipolar Disorder Using Smartphones - the MONARCA iTrial: A Randomized, Placebo-Controlled, Single-Blind, Parallel Group Trial," *Psychological Medicine* (45:13), pp. 2691-2704 (doi: 10.1017/S0033291715000410).

Faurholt-Jepsen, M., Munkholm, K., Frost, M., Bardram, J. E., and Kessing, L. V. 2016. "Electronic Self-Monitoring of Mood Using It Platforms in Adult Patients with

Bipolar Disorder: A Systematic Review of the Validity and Evidence," *BMC Psychiatry* (16:7), pp. 1-14 (doi: 10.1186/s12888-016-0713-0).

*Faurholt-Jepsen, M., Vinberg, M., Frost, M., Christensen, E. M., Bardram, J. E., and Kessing, L. V. 2015. "Smartphone Data as an Electronic Biomarker of Illness Activity in Bipolar Disorder," *Bipolar Disorders* (17:7), pp. 715-728 (doi: 10.1111/bdi.12332).

Feldman, M. S., and Orlikowski, W. J. 2011. "Theorizing Practice and Practicing Theory," Organization Science (22:5), pp. 1240-1253.

*Felipe, S., Singh, A., Bradley, C., Williams, A. C., and Bianchi-Berthouze, N. 2015. "Roles for Personal Informatics in Chronic Pain," in *Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare*. Istanbul, Turkey: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), pp. 161-168 (doi: 10.4108/icst.pervasivehealth.2015.259501).

*Festersen, P. L., and Corradini, A. 2014. "Re: Mind: A Mobile Application for Bipolar Disorder Patients," in *Proceedings of the 4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH 2014)*, Athens, Greece: IEEE, pp. 343-346 (doi:10.4108/icst.mobihealth.2014.257188).

Fine, L. J., Philogene, G. S., Gramling, R., Coups, E. J., and Sinha, S. 2004. "Prevalence of Multiple Chronic Disease Risk Factors: 2001 National Health Interview Survey," *American Journal of Preventive Medicine* (27:2 Supp), pp. 18-24 (doi: 10.1016/j.amepre.2004.04.017).

Finnerty, T., Reeves, S., Dabinett, J., Jeanes, Y. M., and Vögele, C. 2010. "Effects of Peer Influence on Dietary Intake and Physical Activity in Schoolchildren," *Public Health Nutrition* (13:3), pp. 376-383 (doi: 10.1017/S1368980009991315).

Foucault, M. 1980. *Power/Knowledge: Selected Interviews and Other Writings, 1972-1977*, New York, NY: Pantheon Books.

Foucault, M. 1982. "The Subject and Power," *Critical Inquiry* (8:4), pp. 777-795 (https://doi.org/10.1086/448181).

Frantsve, L. M. E., and Kerns, R. D. 2007. "Patient–Provider Interactions in the Management of Chronic Pain: Current Findings within the Context of Shared Medical Decision Making," *Pain Medicine* (8:1), pp. 25-35 (doi: 10.1111/j.1526-4637.2007.00250.x).

*Fukuoka, Y., Kamitani, E., Bonnet, K., and Lindgren, T. 2011. "Real-Time Social Support through a Mobile Virtual Community to Improve Healthy Behavior in Overweight and Sedentary Adults: A Focus Group Analysis," *Journal of Medical Internet Research* (13:3), p. e49 (doi: 10.2196/jmir.1770).

*Fuller, N. R., Fong, M., Gerofi, J., Ferkh, F., Leung, C., Leung, L. S., Zhang, S. Y., Skilton, M., and Caterson, I. D. 2017. "Comparison of an Electronic Versus Traditional Food Diary for Assessing Dietary Intake-a Validation Study," *Obesity Research & Clinical Practice* (11:6), pp. 647-654 (doi: 10.1016/j.orcp.2017.04.001).

Gabriels, K., and Moerenhout, T. 2018. "Exploring Entertainment Medicine and Professionalization of Self-Care: Interview Study among Doctors on the Potential Effects of Digital Self-Tracking," *Journal of medical Internet research* (20:1).

*Garg, S. K., Shah, V. N., Akturk, H. K., Beatson, C., and Snell-Bergeon, J. K. 2017. "Role of Mobile Technology to Improve Diabetes Care in Adults with Type 1 Diabetes: The Remote-T1d Study Ibgstar (R) in Type 1 Diabetes Management," *Diabetes Therapy* (8:4), pp. 811-819.

Gartner. 2018. "Wearables Hold the Key to Connected Health Monitoring." Retrieved Feburary 19, 2019, from https://www.gartner.com/smarterwithgartner/wearables-hold-the-key-to-connected-health-monitoring/

*Gell, N., Grover, K., Humble, M., Sexton, M., Dittus, K., Gell, N. M., and Grover, K. W. 2017. "Efficacy, Feasibility, and Acceptability of a Novel Technology-Based Intervention to Support Physical Activity in Cancer Survivors," *Supportive Care in Cancer* (25:4), pp. 1291-1300 (doi: 10.1007/s00520-016-3523-5).

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*, Boston, MA: Houghton Mifflin.

*Glasgow, R. E., Christiansen, S. M., Kurz, D., King, D. K., Woolley, T., Faber, A. J., Estabrooks, P. A., Strycker, L., Toobert, D., and Dickman, J. 2011. "Engagement in a Diabetes Self-Management Website: Usage Patterns and Generalizability of Program Use," *Journal of Medical Internet Research* (13:1), p. e9 (doi: 10.2196/jmir.1391).

*Goffinet, L., Barrea, T., Beauloye, V., and Lysy, P. A. 2017. "Blood Versus Urine Ketone Monitoring in a Pediatric Cohort of Patients with Type 1 Diabetes: A Crossover Study," *Therapeutic Advances in Endocrinology and Metabolism* (8:1-2), pp. 3-13 (doi: 10.1177/2042018816681706).

Gonzalez, J. S., Shreck, E., Psaros, C., and Safren, S. A. 2015. "Distress and Type 2 Diabetes-Treatment Adherence: A Mediating Role for Perceived Control," *Health Psychology* (34:5), pp. 505-513 (doi: 10.1037/hea0000131).

*Goto, M., Takedani, H., Haga, N., Kubota, M., Ishiyama, M., Ito, S., and Nitta, O. 2014. "Self-Monitoring Has Potential for Home Exercise Programmes in Patients with Haemophilia," *Haemophilia* (20:2), pp. e121-e127 (doi: 10.1111/hae.12355).

Greenlees, I., Graydon, J., and Maynard, I. 2000. "The Impact of Individual Efficacy Beliefs on Group Goal Selection and Group Goal Commitment," *Journal of Sports Sciences* (18:6), pp. 451-459 (doi: 10.1080/02640410050074386).

*Greenwood, D. A., Blozis, S. A., Young, H. M., Nesbitt, T. S., and Quinn, C. C. 2015. "Overcoming Clinical Inertia: A Randomized Clinical Trial of a Telehealth Remote Monitoring Intervention Using Paired Glucose Testing in Adults with Type 2 Diabetes," *Journal of Medical Internet Research* (17:7), p. e178 (doi: 10.2196/jmir.4112).

*Grönvall, E., and Verdezoto, N. 2013. "Beyond Self-Monitoring: Understanding Non-Functional Aspects of Home-Based Healthcare Technology," in *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Zurich, Switzerland: ACM, pp. 587-596 (doi: 10.1145/2493432.2493495).

Grönvall, E., and Verdezoto, N. 2013b. "Understanding Challenges and Opportunities of Preventive Blood Pressure Self-Monitoring at Home," in *Proceedings of the 31st European Conference on Cognitive Ergonomics*. Toulouse, France: ACM, Art. # 31 (doi: 10.1145/2501907.2501962).

*Gu, W., Zhou, Y., Zhou, Z., Liu, X., Zou, H., Zhang, P., Spanos, C. J., and Zhang, L. 2017. "SugarMate: Non-Intrusive Blood Glucose Monitoring with Smartphones," in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies Interact*. (1:3), Art. # 54 (doi: 10.1145/3130919).

*Haak, T., Hanaire, H., Ajjan, R., Hermanns, N., Riveline, J. P., and Rayman, G. 2017. "Use of Flash Glucose-Sensing Technology for 12 Months as a Replacement for Blood Glucose Monitoring in Insulin-Treated Type 2 Diabetes," *Diabetes Therapy: Research, Treatment and Education of Diabetes and Related Disorders* (8:3), pp. 573-586 (doi: 10.1007/s13300-017-0255-6).

*Hales, S., Turner-McGrievy, G. M., Wilcox, S., Davis, R. E., Fahim, A., Huhns, M., and Valafar, H. 2017. "Trading Pounds for Points: Engagement and Weight Loss in a Mobile Health Intervention," *Digital Health* (3) (doi: 10.1177/2055207617702252).

*Hall, S., and Murchie, P. 2014. "Can We Use Technology to Encourage Self-Monitoring by People Treated for Melanoma? A Qualitative Exploration of the Perceptions of Potential Recipients," *Supportive Care in Cancer* (22:6), pp. 1663-1671 (doi: 10.1007/s00520-014-2133-3).

Hamine, S., Gerth-Guyette, E., Faulx, D., Green, B. B., and Ginsburg, A. S. 2015. "Impact of mHealth Chronic Disease Management on Treatment Adherence and Patient Outcomes: A Systematic Review," *Journal of Medical Internet Research* (17:2), p. e52 (doi: 10.2196/jmir.3951).

*Hansen, C. R., Perrild, H., Koefoed, B. G., and Zander, M. 2017. "Video Consultations as Add-on to Standard Care among Patients with Type 2 Diabetes Not Responding to Standard Regimens: A Randomized Controlled Trial," *European Journal of Endocrinology* (176:6), pp. 727-736 (doi: 10.1530/EJE-16-0811).

Heinemann, L., and DeVries, J. H. 2016. "Reimbursement for Continuous Glucose Monitoring," *Diabetes Technology & Therapeutics* (18: Suppl 2), pp. S2-48-S2-52 (doi: 10.1089/dia.2015.0296).

*Hermansen-Kobulnicky, C. J., and Purtzer, M. A. 2014. "Tracking and Journaling the Cancer Journey," *Clinical Journal of Oncology Nursing* (18:4), pp. 388-391 (doi: 10.1188/14.CJON.388-391).

*Hinnen, D. A., Buskirk, A., Lyden, M., Amstutz, L., Hunter, T., Parkin, C. G., and Wagner, R. 2015. "Use of Diabetes Data Management Software Reports by Health Care Providers, Patients with Diabetes, and Caregivers Improves Accuracy and Efficiency of Data Analysis and Interpretation Compared with Traditional Logbook Data: First Results of the Accu-Chek Connect Reports Utility and Efficiency Study (ACCRUES)," *Journal of Diabetes Science and Technology* (9:2), pp. 293-301 (doi: 10.1177/1932296814557188).

Hirschi, T. 2017. *Causes of Delinquency*, New York: NY, Routledge.

Hollenbeck, J. R., Williams, C. R., and Klein, H. J. 1989. "An Empirical Examination of the Antecedents of Commitment to Difficult Goals," *Journal of Applied Psychology* (74:1), pp. 18-23 (https://psycnet.apa.org/doi/10.1037/0021-9010.74.1.18).

Holt, D. T., Helfrich, C. D., Hall, C. G., and Weiner, B. J. 2010. "Are You Ready? How Health Professionals Can Comprehensively Conceptualize Readiness for Change," *Journal of General Internal Medicine* (25: Supp 1), pp. 50-55. 25 (doi: 10.1007/s11606-009-1112-8).

*Hostler, J. M., Sheikh, K. L., Andrada, T. F., Khramtsov, A., Holley, P. R., and Holley, A. B. 2017. "A Mobile, Web-Based System Can Improve Positive Airway Pressure Adherence," *Journal of Sleep Research* (26:2), pp. 139-146 (doi: 10.1111/jsr.12476).

*Hutchesson, M. J., Rollo, M. E., Callister, R., and Collins, C. E. 2015. "Self-Monitoring of Dietary Intake by Young Women: Online Food Records Completed on Computer or Smartphone Are as Accurate as Paper-Based Food Records but More Acceptable," *Journal of the Academy of Nutrition and Dietetics* (115:1), pp. 87-94 (doi: 10.1016/j.jand.2014.07.036).

Huygens, M. W. J., Swinkels, I. C. S., de Jong, J. D., Heijmans, M. J. W. M., Friele, R. D., van Schayck, O. C. P., and de Witte, L. P. 2017. "Self-Monitoring of Health Data by Patients with a Chronic Disease: Does Disease Controllability Matter?," *BMC Family Practice* (18:1), p. 40 (doi: 10.1186/s12875-017-0615-3).

Ilgen, M. A., McKellar, J., Moos, R., and Finney, J. W. 2006. "Therapeutic Alliance and the Relationship between Motivation and Treatment Outcomes in Patients with Alcohol Use Disorder," *Journal of Substance Abuse Treatment* (31:2), pp. 157-162 (doi: 10.1016/j.jsat.2006.04.001).

*Iljaz, R., Brodnik, A., Zrimec, T., and Cukjati, I. 2017. "E-Healthcare for Diabetes Mellitus Type 2 Patients - a Randomised Controlled Trial in Slovenia," *Zdravstveno Varstvo* (56:3), pp. 150-157 (doi: 10.1515/sjph-2017-0020).

*Irace, C., Schweitzer, M. A., Tripolino, C., Scavelli, F. B., and Gnasso, A. 2017. "Diabetes Data Management System to Improve Glycemic Control in People with Type 1 Diabetes: Prospective Cohort Study," *JMIR mHealth and uHealth* (5:11), p. e170 (doi: 10.2196/mhealth.8532).

*Isetta, V., Torres, M., Gonzalez, K., Ruiz, C., Dalmases, M., Embid, C., Navajas, D., Farre, R., and Montserrat, J. M. 2017. "A New mHealth Application to Support Treatment of Sleep Apnoea Patients," *Journal of Telemedicine and Telecare* (23:1), pp. 14-18 (doi: 10.1177/1357633X15621848).

*Izawa, K. P., Watanabe, S., Oka, K., Osada, N., and Omiya, K. 2006. "Effect of Self-Monitoring Approach during Cardiac Rehabilitation on Exercise Maintenance, Self-Efficacy, and Physical Activity over a 1-Year Period after Myocarcial Infarction," *Japanese Journal of Physical Fitness and Sports Medicine* (55: Suppl), pp. S113-S118 (https://doi.org/10.7600/jspfsm.55.S113).

Jakicic, J. M., Davis, K. K., Rogers, R. J, King, W. C., Marcus, M. D., Helsel, D., Rickman, A. D., Wahed, A. S., and Belle, S. H. 2017. "Effect of Wearable Technology Combined with a Lifestyle Intervention on Long-Term Weight Loss: The Idea Randomized Clinical Trial," *JAMA* (316:11), pp. 1161-1171 (doi: 10.1001/jama.2016.12858).

James, T. L., Wallace, L., and Deane, J. K. 2019. "Using Organismic Integration Theory to Explore the Associations between Users' Exercise Motivations and Fitness Technology Feature Set Use," *MIS Quarterly* (43:1), pp. 287-312 (doi: 10.25300/MISQ/2019/14128).

*Ji, L., Su, Q., Feng, B., Shan, Z., Hu, R., Xing, X., Xue, Y., Yang, T., and Hua, Y. 2017. "Structured Self-Monitoring of Blood Glucose Regimens Improve Glycemic Control in Poorly Controlled Chinese Patients on Insulin Therapy: Results from Compass," *Journal of Diabete*s (9:5), pp. 495-501 (doi: 10.1111/1753-0407.12434).

*Johnston, S. K., Nguyen, H. Q., and Wolpin, S. 2009. "Designing and Testing a Web-Based Interface for Self-Monitoring of Exercise and Symptoms for Older Adults with Chronic Obstructive Pulmonary Disease," *Computers, Informatics, Nursing: CIN* (27:3), pp. 166-174 (doi: 10.1097/NCN.0b013e31819f7c1d).

*Jones, M., Lynch, K. T., Kass, A. E., Burrows, A., Williams, J., Wilfley, D. E., and Taylor, C. 2014. "Healthy Weight Regulation and Eating Disorder Prevention in High School Students: A Universal and Targeted Web-Based Intervention," *Journal of Medical Internet Research* (16:2), p. e57 (doi: 10.2196/jmir.2995. 28-39).

*Jongen, P. J., Sinnige, L. G., van Geel, B. M., Verheul, F., Verhagen, W. I., van der Kruijk, R. A., Haverkamp, R., Schrijver, H. M., Baart, J. C., Visser, L. H., Arnoldus, E. P., Gilhuis, H. J., Pop, P., Booy, M., Lemmens, W., Donders, R., Kool, A., and van Noort, E. 2015. "The Interactive Web-based Program MSmonitor for Self-Management and Multidisciplinary Care in Multiple Sclerosis: Utilization and Valuation by Patients," *Patient Preference and Adherence* (10), pp. 243-250 (doi: 10.2147/PPA.S93786).

*Jospe, M. R., Roy, M., Brown, R. C., Williams, S. M., Osborne, H. R., Meredith-Jones, K. A., McArthur, J. R., Fleming, E. A., and Taylor, R. W. 2017. "The Effect of Different Types of Monitoring Strategies on Weight Loss: A Randomized Controlled Trial," *Obesity* (25:9), pp. 1490-1498 (doi: 10.1002/oby.21898).

*Jospe, M. R., Taylor, R. W., Athens, J., Roy, M., and Brown, R. C. 2017. "Adherence to Hunger Training over 6 Months and the Effect on Weight and Eating Behaviour: Secondary Analysis of a Randomised Controlled Trial," *Nutrients* (9:11), p. e1260 (doi: 10.3390/nu9111260).

Kalis, B., Collier, M., and Fu, R. 2018. "Technology: 10 Promising Ai Applications in Health Care," *Harvard Business Review*.

*Karhula, T., Vuorinen, A.-L., Raapysjarvi, K., Pakanen, M., Itkonen, P., Tepponen, M., Junno, U.-M., Jokinen, T., van Gils, M., Lahteenmaki, J., Kohtamaki, K., and Saranummi, N. 2015. "Telemonitoring and Mobile Phone-Based Health Coaching among Finnish Diabetic and Heart Disease Patients: Randomized Controlled Trial," *Journal of Medical Internet Research* (17:6), p. e153 (doi: 10.2196/jmir.4059).

*Kempf, K., Altpeter, B., Berger, J., Reuβ, O., Fuchs, M., Schneider, M., Gärtner, B., Niedermeier, K., and Martin, S. 2017. "Efficacy of the Telemedical Lifestyle Intervention Program Telipro in Advanced Stages of Type 2 Diabetes: A Randomized Controlled Trial," *Diabetes Care* (40:7), pp. 863-871 (doi: 10.2337/dc17-0303).

*Kendall, L., Morris, D., and Tan, D. 2015. "Blood Pressure Beyond the Clinic: Rethinking a Health Metric for Everyone," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Seoul, Republic of Korea: ACM, pp. 1679-1688 (doi: 10.1145/2702123.2702276).

Kennedy, C. M., Powell, J., Payne, T. H., Ainsworth, J., Boyd, A., and Buchan, I. 2012. "Active Assistance Technology for Health-Related Behavior Change: An Interdisciplinary Review," *Journal of Medical Internet Research* (14:3), p. e80 (doi: 10.2196/jmir.1893).

Klein, H. J., Wesson, M. J., Hollenbeck, J. R., and Alge, B. J. 1999. "Goal Commitment and the Goal-Setting Process: Conceptual Clarification and Empirical Synthesis," *Journal of Applied Psychology* (84:6), p. 885-896 (https://psycnet.apa.org/doi/10.1037/0021-9010.84.6.885).

*Kolodziejczyk, J. K., Norman, G. J., Rock, C. L., Arredondo, E. M., Madanat, H., Roesch, S. C., and Patrick, K. 2014. "Strategies That Predict Weight Loss among

Overweight/Obese Young Adults," *American Journal of Health Behavior* (38:6), pp. 871-880 (doi: 10.5993/AJHB.38.6.9).

*Krukowski, R. A., Harvey-Berino, J., Bursac, Z., Ashikaga, T., and West, D. S. 2013. "Patterns of Success: Online Self-Monitoring in a Web-Based Behavioral Weight Control Program," *Health Psychology* (32:2), pp. 164-170 (doi: 10.1037/a0028135).

*Laing, B. Y., Mangione, C. M., Chi-Hong, T., Mei, L., Vaisberg, E., Mahida, M., Bholat, M., Glazier, E., Morisky, D. E., and Bell, D. S. 2015. "Effectiveness of a Smartphone Application for Weight Loss Compared with Usual Care in Overweight Primary Care Patients," *Annals of Internal Medicine* (161:10 Suppl), p. S5 (doi: 10.7326/M13-30050.

*Langstrup, H., and Winthereik, B. R. 2008. "The Making of Self-Monitoring Asthma Patients: Mending a Split Reality with Comparative Ethnography," *Comparative Sociology* (7:3), pp. 362-386 (doi: 10.1163/156913308X306663).

Leary, M. R., and Baumeister, R. F. 2017. "The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation," in: *Interpersonal Development*. Routledge, pp. 57-89.

*Lee, M. K., Lee, K. H., Yoo, S. H., and Park, C. Y. 2017. "Impact of Initial Active Engagement in Self-Monitoring with a Telemonitoring Device on Glycemic Control among Patients with Type 2 Diabetes," *Scientific Reports* (7:1), p. 3866 (doi: 10.1038/s41598-017-03842-2).

Legorreta, A. P., Leung, K.-M., Berkbigler, D., Evans, R., and Liu, X. 2000. "Outcomes of a Population-Based Asthma Management Program: Quality of Life, Absenteeism, and Utilization," *Annals of Allergy, Asthma & Immunology* (85:1), pp. 28-34 (doi: 10.1016/S1081-1206(10)62430-1).

Lehrig, T., Krancher, O., and Dibbern, J. 2017. "How Users Perceive and Actualize Affordances: An Exploratory Case Study of Collaboration Platforms," in *Proceedings of Thirty Eighth International Conference on Information Systems*, Seoul, South Korea.

Lehto, T., and Oinas-Kukkonen, H. 2011. "Persuasive Features in Web-Based Alcohol and Smoking Interventions: A Systematic Review of the Literature," *Journal of Medical Internet Research* (13:3), p. e46 (doi: 10.2196/jmir.1559).

Leinonen, A. M., Pyky, R., Ahola, R., Kangas, M., Siirtola, P., Luoto, T., Enwald, H., Iakheimo, T. M., Roning, J., Keinanen-Kiukaanniemi, S., Mantysaari, M., Korpelainen, R., and Jamsa, T. 2017. "Feasibility of Gamified Mobile Service Aimed at Physical Activation in Young Men: Population-Based Randomized Controlled Study (MOPO)," *JMIR mHealth and uHealth* (5:10), p. e146 (doi: 10.2196/mhealth.6675).

Li, I., Dey, A., and Forlizzi, J. 2010. "A Stage-Based Model of Personal Informatics Systems," in *Proceedings of the SIGCHI Conference on Human Factors in*

Computing Systems. Atlanta, Georgia, USA: ACM, pp. 557-566 (doi: 10.1145/1753326.1753409).

Liang, Z. L., Liu, W. Y., Ploderer, B., Bailey, J., Kulik, L., and Li, Y. X. 2017. "Designing Intelligent Sleep Analysis Systems for Automated Contextual Exploration on Personal Sleep-Tracking Data," in *New Frontiers in Artificial Intelligence*, M. Otake, et al. (eds.), Kanagawa, Japan: Springer International Publishing, pp. 367-379.

Locke, E. A. 1991. "Goal Theory vs. Control Theory: Contrasting Approaches to Understanding Work Motivation," *Motivation and Emotion* (15:1), pp. 9-28 (doi: 10.1007/BF00991473).

Locke, E. A., and Latham, G. P. 2002. "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey," *American Psychologist* (57:9), pp. 705-717 (https://psycnet.apa.org/doi/10.1037/0003-066X.57.9.705).

Lorig, K. 1996. "Chronic Disease Self-Management: A Model for Tertiary Prevention," *American Behavioral Scientist* (39:6), pp. 676-683 (doi: 10.1177/0002764296039006005).

Lorig, K. R., Sobel, D. S., Stewart, A. L., Brown Jr, B. W., Bandura, A., Ritter, P., Gonzalez, V. M., Laurent, D. D., and Holman, H. R. 1999. "Evidence Suggesting That a Chronic Disease Self-Management Program Can Improve Health Status While Reducing Hospitalization: A Randomized Trial," *Medical Care* (37:1), pp. 5-14 (doi: 10.1097/00005650-199901000-00003).

Lupton, D. 2014. "Self-Tracking Cultures: Towards a Sociology of Personal Informatics," in *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design. Sydney*. New South Wales, Australia: ACM, pp. 77-86 (doi: 10.1145/2686612.2686623).

Luszczynska, A., Gibbons, F. X., Piko, B. F., and Tekozel, M. 2004. "Self-Regulatory Cognitions, Social Comparison, and Perceived Peers' Behaviors as Predictors of Nutrition and Physical Activity: A Comparison among Adolescents in Hungary, Poland, Turkey, and USA," *Psychology & Health* (19:5), pp. 577-593 (doi: 10.1080/0887044042000205844).

*Ma, J., Xiao, L., and Blonstein, A. C. 2013. "Measurement of Self-Monitoring Web Technology Acceptance and Use in an e-Health Weight-Loss Trial," *Telemedicine and e-Health* (19:10), pp. 739-745 (doi: 10.1089/tmj.2013.0009).

Maddux, J. E., and Rogers, R. W. 1983. "Protection Motivation and Self-Efficacy: A Revised Theory of Fear Appeals and Attitude Change," *Journal of Experimental Social Psychology* (19:5), pp. 469-479 (doi: 10.1016/0022-1031(83)90023-9).

Madrian, B. C., and Shea, D. F. 2001. "The Power of Suggestion: Inertia in 401 (K) Participation and Savings Behavior," *The Quarterly journal of economics* (116:4), pp. 1149-1187.

Mallery, L., and Rockwood, K. 1992. "Preventive Care for the Elderly: Uncovering the Unmet Needs of This Population," *Canadian Family Physician* (38), pp. 2371-2379.

*Mantani, A., Kato, T., Furukawa, T. A., Horikoshi, M., Imai, H., Hiroe, T., Chino, B., Funayama, T., Yonemoto, N., Zhou, Q., and Kawanishi, N. 2017. "Smartphone Cognitive Behavioral Therapy as an Adjunct to Pharmacotherapy for Refractory Depression: Randomized Controlled Trial," *Journal of Medical Internet Research* (19:11), p. e373 (doi: 10.2196/jmir.8602).

Martin, C. M. 2007. "Chronic Disease and Illness Care: Adding Principles of Family Medicine to Address Ongoing Health System Redesign," *Canadian Family Physician* (53:12), pp. 2086-2091.

Marx, G. T. 2002. "What's New About the" New Surveillance"? Classifying for Change and Continuity," *Surveillance & Society* (1:1), pp. 9-29 (doi: 10.24908/ss.v1i1.3391).

*Mathieu-Fritz, A., and Guillot, C. 2017. "Diabetes Self-Monitoring Devices and "Patient Work" Transformations. New Forms of Temporality, Reflexivity and Self-Knowledge Related to the Experience of Chronic Disease," *Revue d'Anthropologie des Connaissances* (11:4), pp. K-AO (doi: 10.3917/rac.037.k).

*Matthews, M., Murnane, E., and Snyder, J. 2017. "Quantifying the Changeable Self: The Role of Self-Tracking in Coming to Terms with and Managing Bipolar Disorder," *Human-Computer Interaction* (32:5-6), pp. 413-446 (doi: 10.1080/07370024.2017.1294983).

McBain, H., Shipley, M., and Newman, S. 2015. "The Impact of Self-Monitoring in Chronic Illness on Healthcare Utilisation: A Systematic Review of Reviews," *BMC Health Services Research* (15:565), p. 1-10 (doi: 10.1186/s12913-015-1221-5).

*McDonald, L., Glen, F. C., Taylor, D. J., and Crabb, D. P. 2017. "Self-Monitoring Symptoms in Glaucoma: A Feasibility Study of a Web-Based Diary Tool," *Journal of Ophthalmology*, (2017: Art. ID 8452840) (doi: 10.1155/2017/8452840).

*McKnight, R. F., Bilderbeck, A. C., Miklowitz, D. J., Hinds, C., Goodwin, G. M., and Geddes, J. R. 2017. "Longitudinal Mood Monitoring in Bipolar Disorder: Course of Illness as Revealed through a Short Messaging Service," *Journal of Affective Disorders* (223), pp. 139-145 (doi: 10.1016/j.jad.2017.07.029).

*Mentis, H. M., Komlodi, A., Schrader, K., Phipps, M., Gruber-Baldini, A., Yarbrough, K., and Shulman, L. 2017. "Crafting a View of Self-Tracking Data in the Clinical Visit," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver, Colorado, USA: ACM, pp. 5800-5812 (doi: 10.1145/3025453.3025589).

Mettler, T., and Wulf, J. 2019. "Physiolytics at the Workplace: Affordances and Constraints of Wearables Use from an Employee's Perspective," *Information Systems Journal* (29:1), pp. 245-273 (doi: 10.1111/isj.12205).

Michie, S., van Stralen, M. M., and West, R. 2011. "The Behaviour Change Wheel: A New Method for Characterising and Designing Behaviour Change Interventions," *Implementation Science* (6), p. 42 (doi: 10.1186/1748-5908-6-42).

Milgrom, P., and Roberts, J. 1995. "Complementarities and Fit Strategy, Structure, and Organizational Change in Manufacturing," *Journal of Accounting and Economics* (19:2-3), pp. 179-208 (doi: 10.1016/0165-4101(94)00382-F).

Minet, L., Møller, S., Vach, W., Wagner, L., and Henriksen, J. E. 2010. "Mediating the Effect of Self-Care Management Intervention in Type 2 Diabetes: A Meta-Analysis of 47 Randomised Controlled Trials," *Patient Education and Counseling* (80:1), pp. 29-41 (doi: 10.1016/j.pec.2009.09.033).

*Morgan, P., Scott, H., Young, M., Plotnikoff, R., Collins, C., and Callister, R. 2014. "Associations between Program Outcomes and Adherence to Social Cognitive Theory Tasks: Process Evaluation of the SHED-IT Community Weight Loss Trial for Men," *The International Journal of Behavioral Nutrition & Physical Activity* (11:1), p. 89 (doi: 10.1186/s12966-014-0089-9).

*Mouzouras, N., Pogiatzis, A., Kleanthous, S., and Samaras, G. 2017. "Dermatrack: A Skin Cancer Tracking Intelligent Application," in *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion*. Limassol, Cyprus: ACM, pp. 85-88 (doi: 10.1145/3030024.3038271).

Mueller, A. S., Pearson, J., Muller, C., Frank, K., and Turner, A. 2010. "Sizing up Peers: Adolescent Girls' Weight Control and Social Comparison in the School Context," *Journal of Health and Social Behavior* (51:1), pp. 64-78 (doi: 10.1177/0022146509361191).

*Mummah, S., Robinson, T. N., Mathur, M., Farzinkhou, S., Sutton, S., and Gardner, C. D. 2017. "Effect of a Mobile App Intervention on Vegetable Consumption in Overweight Adults: A Randomized Controlled Trial," *The International Journal of Behavioral Nutrition and Physical Activity* (14: Art. ID 125), p.125 (https://doi.org/10.1186/s12966-017-0563-2).

*Munster-Segev, M., Fuerst, O., Kaplan, S. A., and Cahn, A. 2017. "Incorporation of a Stress Reducing Mobile App in the Care of Patients with Type 2 Diabetes: A Prospective Study," *JMIR mHealth and uHealth* (5), p. e75 (doi: 10.2196/mhealth.7408).

*Murnane, E. L., Cosley, D., Chang, P., Guha, S., Frank, E., Gay, G., and Matthews, M. 2016. "Self-Monitoring Practices, Attitudes, and Needs of Individuals with Bipolar Disorder: Implications for the Design of Technologies to Manage Mental Health," *Journal of the American Medical Informatics Association* (23:3), pp. 477-84 (doi: 10.1093/jamia/ocv165).

*Nakano, H., Kikuya, M., Hara, A., Nakashita, M., Hirose, T., Obara, T., Metoki, H., Inoue, R., Asayama, K., Ohkubo, T., Totsune, K., and Imai, Y. 2011. "Self-Monitoring of Ambulatory Blood Pressure by the Microlife WatchBP O3 – an Application Test," *Clinical & Experimental Hypertension* (33:1), pp. 34-40 (doi: 10.3109/10641963.2010.503300).

Nam, S., Chesla, C., Stotts, N. A., Kroon, L., and Janson, S. L. 2011. "Barriers to Diabetes Management: Patient and Provider Factors," *Diabetes Research and Clinical Practice* (93:1), pp. 1-9 (doi: 10.1016/j.diabres.2011.02.002).

*Naylor, M. R., Keefe, F. J., Brigidi, B., Naud, S., and Helzer, J. E. 2008. "Therapeutic Interactive Voice Response for Chronic Pain Reduction and Relapse Prevention," *Pain* (134:3), pp. 335-345 (doi: 10.1016/j.pain.2007.11.001).

*Nicklas, B. J., Gaukstern, J. E., Beavers, K. M., Newman, J. C., Leng, X. Y., and Rejeski, W. J. 2014. "Self-Monitoring of Spontaneous Physical Activity and Sedentary Behavior to Prevent Weight Regain in Older Adults," *Obesity* (22:6), pp. 1406-1412 (doi: 10.1002/oby.20732).

*Nishimura, A., Harashima, S., Fujita, Y., Tanaka, D., Wang, Y., Liu, Y. Y., and Inagaki, N. 2017. "Effects of Structured Testing Versus Routine Testing of Blood Glucose in Diabetes Self-Management: A Randomized Controlled Trial," *Journal of Diabetes and Its Complications* (31:1), pp. 228-233 (doi: 10.1016/j.jdiacomp.2016.08.019).

*Nørregaard, L. B., Løventoft, P. K., Frøkjær, E., Lauritsen, L., Olsson, E. C., Andersen, L., Rauff, S., and Martiny, K. 2014. "Patient Expectations and Experiences from a Clinical Study in Psychiatric Care Using a Self-Monitoring System," in *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. Helsinki, Finland: ACM, pp. 991-994 (doi: 10.1145/2639189.2670258).

Norris, S. L., Engelgau, M. M., and Narayan, K. V. 2001. "Effectiveness of Self-Management Training in Type 2 Diabetes," *Diabetes Care* (24:3), pp. 561-587 (https://doi.org/10.2337/diacare.24.3.561).

Oinas-Kukkonen, H., and Harjumaa, M. 2009. "Persuasive Systems Design: Key Issues, Process Model, and System Features," *Communications of the Association for Information Systems* (24: Art. 28), p. 485-500 (doi: 10.17705/1CAIS.02428).

Okoli, C., and Schabram, K. 2010. "A Guide to Conducting a Systematic Literature Review of Information Systems Research," *SSRN Working Paper Series* (10:26), 2 pages (doi: 10.2139/ssrn.1954824).

*Olafsdottir, A. F., Attvall, S., Sandgren, U., Dahlqvist, S., Pivodic, A., Skrtic, S., Theodorsson, E., and Lind, M. 2017. "A Clinical Trial of the Accuracy and Treatment Experience of the Flash Glucose Monitor Freestyle Libre in Adults with Type 1 Diabetes," *Diabetes Technology & Therapeutics* (19:3), pp. 164-172 (doi: 10.1089/dia.2016.0392).

*Or, C., and Tao, D. 2016. "A 3-Month Randomized Controlled Pilot Trial of a Patient-Centered, Computer-Based Self-Monitoring System for the Care of Type 2 Diabetes Mellitus and Hypertension," *Journal of Medical Systems* (40:4), p. 81 (doi: 10.1007/s10916-016-0437-1).

*Painter, S. L., Ahmed, R., Hill, J. O., Kushner, R. F., Lindquist, R., Brunning, S., and Margulies, A. 2017. "What Matters in Weight Loss? An in-Depth Analysis of Self-Monitoring," *Journal of Medical Internet Research* (19:5), p. e160 (doi: 10.2196/jmir.7457).

Panagioti, M., Richardson, G., Small, N., Murray, E., Rogers, A., Kennedy, A., Newman, S., and Bower, P. 2014. "Self-Management Support Interventions to Reduce Health Care Utilisation without Compromising Outcomes: A Systematic Review and Meta-Analysis," *BMC Health Services Research* (14), p. 356 (doi: 10.1186/1472-6963-14-356).

*Partridge, S. R., Allman-Farinelli, M., McGeechan, K., Balestracci, K., Wong, A. T. Y., Hebden, L., Harris, M. F., Bauman, A., and Phongsavan, P. 2016. "Process Evaluation of TXT2BFiT: A Multi-Component mHealth Randomised Controlled Trial to Prevent Weight Gain in Young Adults," *International Journal of Behavioral Nutrition and Physical Activity* (13), p. 7 (doi: 10.1186/s12966-016-0329-2).

Paterson, B. L., Russell, C., and Thorne, S. 2001. "Critical Analysis of Everyday Self-Care Decision Making in Chronic Illness," *Journal of Advanced Nursing* (35:3), pp. 335-341.

*Paula, J. S., Braga, L. D., Moreira, R. O., and Kupfer, R. 2017. "Correlation between Parameters of Self-Monitoring of Blood Glucose and the Perception of Health-Related Quality of Life in Patients with Type 1 Diabetes Mellitus," *Archives of Endocrinology Metabolism* (61:4), pp. 343-347 (doi: 10.1590/2359-3997000000222).

*Pedersen, N., Elkjaer, M., Duricova, D., Burisch, J., Dobrzanski, C., Andersen, N. N., Jess, T., Bendtsen, F., Langholz, E., Leotta, S., Knudsen, T., Thorsgaard, N., and Munkholm, P. 2012. "eHealth: Individualisation of Infliximab Treatment and Disease Course via a Self-Managed Web-Based Solution in Crohn's Disease," *Alimentary Pharmacology and Therapeutics* (36:9), pp. 840-849 (doi: 10.1111/apt.12043).

Peytremann-Bridevaux, I., and Burnand, B. 2009. "Disease Management: A Proposal for a New Definition," *International Journal of Integrated Care* (9), p. e16.

*Piras, E. M., and Miele, F. 2017. "Clinical Self-Tracking and Monitoring Technologies: Negotiations in the ICT-Mediated Patient-Provider Relationship," *Health Sociology Review* (26:1), pp. 38-53 (https://doi.org/10.1080/14461242.2016.1212316).

Piwek, L., Ellis, D. A., Andrews, S., and Joinson, A. 2016. "The Rise of Consumer Health Wearables: Promises and Barriers," *PLoS Medicine* (13:2), p. e1001953 (doi: 10.1371/journal.pmed.1001953).

*Plow, M., and Golding, M. 2017. "Using mHealth Technology in a Self-Management Intervention to Promote Physical Activity among Adults with Chronic Disabling Conditions: Randomized Controlled Trial," *JMIR mHealth and uHealth* (5:12), p. e185 (doi: 10.2196/mhealth.6394).

Polites, G. L., and Karahanna, E. 2013. "The Embeddedness of Information Systems Habits in Organizational and Individual Level Routines: Development and Disruption," *MIS Quarterly* (37:1), pp. 221-246.

*Polonsky, W. H., Hessler, D., Ruedy, K. J., Beck, R. W., and Group, D. S. 2017. "The Impact of Continuous Glucose Monitoring on Markers of Quality of Life in Adults with Type 1 Diabetes: Further Findings from the DIAMOND Randomized Clinical Trial," *Diabetes Care* (40:6), pp. 736-741 (doi: 10.2337/dc17-0133).

Prochaska, J. O., and DiClemente, C. C. 1982. "Transtheoretical Therapy: Toward a More Integrative Model of Change," *Psychotherapy: Theory, Research & Practice* (19:3), pp. 276-288 (doi: 10.1037/h0088437).

*Rader, S., Dorner, T. E., Schoberberger, R., and Wolf, H. 2017. "Effects of a Web-Based Follow-up Intervention on Self-Efficacy in Obesity Treatment for Women," *Wiener Klinische Wochenschrift* (129:13-14), pp. 472-481 (doi: 10.1007/s00508-017-1198-7).

*Raiff, B. R., and Dallery, J. 2010. "Internet-Based Contingency Management to Improve Adherence with Blood Glucose Testing Recommendations for Teens with Type 1 Diabetes," *Journal of Applied Behavior Analysis* (43:3), pp. 487-491 (doi: 10.1901/jaba.2010.43-487).

*Ramanathan, N., Swendeman, D., Comulada, W. S., Estrin, D., and Rotheram-Borus, M. J. 2013. "Identifying Preferences for Mobile Health Applications for Self-Monitoring and Self-Management: Focus Group Findings from HIV-Positive Persons and Young Mothers," *International Journal of Medical Informatics* (82:4), p. e38-46 (doi: 10.1016/j.ijmedinf.2012.05.009).

*Roblin, D. W. 2011. "The Potential of Cellular Technology to Mediate Social Networks for Support of Chronic Disease Self-Management," *Journal of Health Communication* (16: Suppl 1), pp. 59-76 (doi: 10.1080/10810730.2011.596610.).

Ruckenstein, M., and Schull, N. D. 2017. "The Datafication of Health," in *Annual Review of Anthropology*, D. Brenneis and K.B. Strier (eds.), pp. 261-278.

*Ruotsalainen, H., Kyngäs, H., Tammelin, T., Heikkinen, H., and Kääriäinen, M. 2015. "Effectiveness of Facebook-Delivered Lifestyle Counselling and Physical Activity Self-Monitoring on Physical Activity and Body Mass Index in Overweight and Obese Adolescents: A Randomized Controlled Trial," *Nursing*

Research and Practice (2015: Art. ID 159205), 14 pages (doi: 10.1155/2015/159205).

Ryan, R. M., and Deci, E. L. 2000. "Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions," *Contemporary Educational Psychology* (25:1), pp. 54-67 (https://doi.org/10.1006/ceps.1999.1020).

*Ryan, D., Price, D., Musgrave, S. D., Malhotra, S., Lee, A. J., Ayansina, D., Sheikh, A., Tarassenko, L., Pagliari, C., and Pinnock, H. 2012. "Clinical and Cost Effectiveness of Mobile Phone Supported Self-Monitoring of Asthma: Multicentre Randomised Controlled Trial," *BMJ: British Medical Journal* (344: Art. e1756), (https://doi.org/10.1136/bmj.e1756).

*Sage, A., Roberts, C., Geryk, L., Sleath, B., Tate, D., and Carpenter, D. 2017. "A Self-Regulation Theory–Based Asthma Management Mobile App for Adolescents: A Usability Assessment," *JMIR Human Factors* (4:1), p. e5 (doi: 10.2196/humanfactors.7133).

Samuelson, P. A. 1974. "Complementarity: An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory," *Journal of Economic literature* (12:4), pp. 1255-1289.

*Sasai, H., Ueda, K., Tsujimoto, T., Kobayashi, H., Sanbongi, C., Ikegami, S., and Nakata, Y. 2017. "Dose-Ranging Pilot Randomized Trial of Amino Acid Mixture Combined with Physical Activity Promotion for Reducing Abdominal Fat in Overweight Adults," *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* (10), pp. 297-309 (doi: 10.2147/DMSO.S138084).

*Scharer, L. O., Krienke, U. J., Graf, S.-M., Meltzer, K., and Langosch, J. M. 2015. "Validation of Life-Charts Documented with the Personal Life-Chart App-a Self-Monitoring Tool for Bipolar Disorder," *BMC Psychiatry* (15), p. 49 (doi: 10.1186/s12888-015-0414-0).

Schilling, L. S., Grey, M., and Knafl, K. A. 2002. "The Concept of Self-Management of Type 1 Diabetes in Children and Adolescents: An Evolutionary Concept Analysis," *Journal of Advanced Nursing* (37:1), pp. 87-99 (doi: 10.1046/j.1365-2648.2002.02061.x).

Schneider, C. Q., and Wagemann, C. 2010. "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets," *Comparative Sociology* (9:3), pp. 397-418 (doi: 10.1163/156913210X12493538729793).

*Schroder, K. E. E. 2011. "Computer-Assisted Dieting: Effects of a Randomized Nutrition Intervention," *American Journal of Health Behavior* (35:2), pp. 175-188.

*Selvan, C., Thukral, A., Dutta, D., Ghosh, S., and Chowdhury, S. 2017. "Impact of Self-Monitoring of Blood Glucose Log Reliability on Long-Term Glycemic Outcomes in Children with Type 1 Diabetes," *Indian Journal of Endocrinology & Metabolism* (21:3), pp. 382-386 (doi: 10.4103/ijem.IJEM_342_16).

*Sevick, M. A., Zickmund, S., Korytkowski, M., Piraino, B., Sereika, S., Mihalko, S., Snetselaar, L., Stumbo, P., Hausmann, L., Ren, D. X., Marsh, R., Sakraida, T., Gibson, J., Safaien, M., Starrett, T. J., and Burke, L. E. 2008. "Design, Feasibility, and Acceptability of an Intervention Using Personal Digital Assistant-Based Self-Monitoring in Managing Type 2 Diabetes," *Contemporary Clinical Trials* (29:3), pp. 396-409 (https://doi.org/10.1016/j.cct.2007.09.004).

*Sevick, M., Stone, R., Zickmund, S., Wang, Y., Korytkowski, M., and Burke, L. 2010. "Factors Associated with Probability of Personal Digital Assistant-Based Dietary Self-Monitoring in Those with Type 2 Diabetes," *Journal of Behavioral Medicine* (33:4), pp. 315-325 (doi: 10.1007/s10865-010-9257-9).

*Shaiful, A. K. M., Manaf, M. R. A., Nor, N. S. M., and Ambak, R. 2017. "Effects of Lifestyle Intervention towards Obesity and Blood Pressure among Housewives in Klang Valley: A Quasi-Experimental Study," *The Malaysian Journal of Medical Sciences* (24:6), pp. 83-91 (doi: 10.21315/mjms2017.24.6.10).

Sharon, T., and Zandbergen, D. 2017. "From Data Fetishism to Quantifying Selves: Self-Tracking Practices and the Other Values of Data," *New Media & Society* (19:11), pp. 1695-1709 (doi: 10.1177/1461444816636090).

Shin, D. H., and Biocca, F. 2017. "Health Experience Model of Personal Informatics: The Case of a Quantified Self," *Computers in Human Behavior* (69), pp. 62-74 (https://doi.org/10.1016/j.chb.2016.12.019).

*Shuger, S. L., Barry, V. W., Xuemei, S., McClain, A., Hand, G. A., Wilcox, S., Meriwether, R. A., Hardin, J. W., and Blair, S. N. 2011. "Electronic Feedback in a Diet- and Physical Activity-Based Lifestyle Intervention for Weight Loss: A Randomized Controlled Trial," *The International Journal of Behavioral Nutrition and Physical Activity* (8), p. 41 (doi: 10.1186/1479-5868-8-41).

*Sidhu, M. S., Daley, A., and Jolly, K. 2016. "Evaluation of a Text Supported Weight Maintenance Programme 'Lighten up Plus' Following a Weight Reduction Programme: Randomised Controlled Trial," *The International Journal of Behavioral Nutrition & Physical Activity* (13), p. 19 (doi: 10.1186/s12966-016-0346-1).

*Sieber, J., Flacke, F., Link, M., Haug, C., and Freckmann, G. 2017. "Improved Glycemic Control in a Patient Group Performing 7-Point Profile Self-Monitoring of Blood Glucose and Intensive Data Documentation: An Open-Label, Multicenter, Observational Study," *Diabetes Therapy* (8:5), pp. 1079-1085.

*Simons, C. J. P., Drukker, M., Evers, S., van Mastrigt, G., Hohn, P., Kramer, I., Peeters, F., Delespaul, P., Menne-Lothmann, C., Hartmann, J. A., van Os, J., and Wichers, M. 2017. "Economic Evaluation of an Experience Sampling Method Intervention in Depression Compared with Treatment as Usual Using Data from a Randomized Controlled Trial," *BMC Psychiatry* (17:1), p. 415 (doi: 10.1186/s12888-017-1577-7).

Sun, H., Fang, Y., and Zou, H. M. 2016. "Choosing a Fit Technology: Understanding Mindfulness in Technology Adoption and Continuance," *Journal of the Association for Information Systems* (17:6), pp. 377-412 (doi: 10.17705/1jais.00431).

*Spring, B., Pellegrini, C. A., Pfammatter, A., Duncan, J. M., Pictor, A., McFadden, H. G., Siddique, J., and Hedeker, D. 2017. "Effects of an Abbreviated Obesity Intervention Supported by Mobile Technology: The ENGAGED Randomized Clinical Trial," *Obesity* (25:7), pp. 1191-1198 (doi: 10.1002/oby.21842).

Standage, M., Sebire, S. J., and Loney, T. 2008. "Does Exercise Motivation Predict Engagement in Objectively Assessed Bouts of Moderate-Intensity Exercise?: A Self-Determination Theory Perspective," *Journal of Sport and Exercise Psychology* (30:4), pp. 337-352 (doi: 10.1123/jsep.30.4.337).

*Stark, S., Snetselaar, L., Piraino, B., Stone, R. A., Kim, S., Hall, B., Burke, L. E., and Sevick, M. A. 2011. "Personal Digital Assistant-Based Self-Monitoring Adherence Rates in 2 Dialysis Dietary Intervention Pilot Studies: BalanceWise-HD and BalanceWise-PD," *Journal of Renal Nutrition* (21:6), pp. 492-498 (doi: 10.1053/j.jrn.2010.10.026).

Stein, M.-K., Newell, S., Wagner, E. L., and Galliers, R. D. 2015. "Coping with Information Technology: Mixed Emotions, Vacillation, and Nonconforming Use Patterns," *MIS Quarterly* (39:2), pp. 367-392 (doi: 10.25300/MISQ/2015/39.2.05).

*Steinberg, D. M., Christy, J., Batch, B. C., Askew, S., Moore, R. H., Parker, P., and Bennett, G. G. 2017. "Preventing Weight Gain Improves Sleep Quality among Black Women: Results from a RCT," *Annals of Behavioral Medicine* (51:4), pp. 555-566 (doi: 10.1007/s12160-017-9879-z).

*Steinberg, D. M., Levine, E. L., Askew, S., Foley, P., and Bennett, G. G. 2013. "Daily Text Messaging for Weight Control among Racial and Ethnic Minority Women: Randomized Controlled Pilot Study," *Journal of Medical Internet Research* (15:11), p. e244 (doi: 10.2196/jmir.2844).

Steinberg, D. M., Levine, E. L., Lane, I., Askew, S., Foley, P. B., Puleo, E., and Bennett, G. G. 2014. "Adherence to Self-Monitoring via Interactive Voice Response Technology in an eHealth Intervention Targeting Weight Gain Prevention among Black Women: Randomized Controlled Trial," *Journal of Medical Internet Research* (16:4), p. e114 (doi: 10.2196/jmir.2996).

*Storni, C. 2010. "Multiple Forms of Appropriation in Self-Monitoring Technology: Reflections on the Role of Evaluation in Future Self-Care," *International Journal of Human-Computer Interaction* (26:5), pp. 537-561 (doi: 10.1080/10447311003720001).

*Storni, C. 2014. "Design Challenges for Ubiquitous and Personal Computing in Chronic Disease Care and Patient Empowerment: A Case Study Rethinking

Diabetes Self-Monitoring," *Personal and Ubiquitous Computing* (18:5), pp. 1277-1290 (doi: 10.1007/s00779-013-0707-6).

*Storni, C. 2014. "Diabetes Self-Care in-the-Wild: Design Challenges for Personal Health Record Systems and Self-Monitoring Technologies," *Information Technology & People* (27:4), pp. 397-420 (doi: 10.1108/ITP-02-2013-0032).

Strack, F., and Deutsch, R. 2011. "A Theory of Impulse and Reflection," *Handbook of Theories of Social Psychology*, pp. 97-117 (doi: 10.4135/9781446249215.n6).

Strong, D. M., Johnson, S. A., Tulu, B., Trudel, J., Volkoff, O., Pelletier, L. R., Bar-On, I., and Garber, L. 2014. "A Theory of Organization-EHR Affordance Actualization," *Journal of the Association for Information Systems* (15:2), pp. 53-85 (doi: 10.17705/1jais.00353).

Suh, A. 2018. "Sustaining the Use of Quantified-Self Technology: A Theoretical Extension and Empirical Test," *Asia Pacific Journal of Information Systems* (28:2), pp. 114-132 (doi: 10.14329/apjis.2018.28.2.114).

Swan, M. 2009. "Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking," *International Journal of Environmental Research and Public Health* (6:2), pp. 492-525 (doi: 10.3390/ijerph6020492).

Swann, W. B., Pelham, B. W., and Krull, D. S. 1989. "Agreeable Fancy or Disagreeable Truth? Reconciling Self-Enhancement and Self-Verification," *Journal of Personality and Social Psychology* (57:5), pp. 782-791 (doi:10.1037/0022-3514.57.5.782).

*Swendeman, D., Ramanathan, N., Baetscher, L., Medich, M., Scheffler, A., Comulada, W., and Estrin, D. 2015. "Smartphone Self-Monitoring to Support Self-Management among People Living with HIV: Perceived Benefits and Theory of Change from a Mixed-Methods Randomized Pilot Study," *Journal of Acquired Immune Deficiency Syndromes* (69: Suppl 1), pp. S80-S91 (doi: 10.1097/QAI.0000000000000570).

Tarafdar, M., Cooper, C. L., and Stich, J. F. 2019. "The Technostress Trifecta-Techno Eustress, Techno Distress and Design: Theoretical Directions and an Agenda for Research," *Information Systems Journal* (29:1), pp. 6-42 (doi: 10.1111/isj.12169/full).

Tay, I., Garland, S., Gorelik, A., and Wark, J. D. 2017. "Development and Testing of a Mobile Phone App for Self-Monitoring of Calcium Intake in Young Women," *JMIR mHealth and uHealth* (5:3), p. e27 (doi: 10.2196/mhealth.5717).

Thapa, D., and Sein, M. K. 2018. "Trajectory of Affordances: Insights from a Case of Telemedicine in Nepal," *Information Systems Journal* (28:5), pp. 796-817 (https://doi.org/10.1111/isj.12160).

*Thomas, J. G., Leahey, T. M., and Wing, R. R. 2015. "An Automated Internet Behavioral Weight-Loss Program by Physician Referral: A Randomized Controlled Trial," *Diabetes Care* (38:1), pp. 9-15 (doi: 10.2337/dc14-1474).

Thompson, J. K., Shroff, H., Herbozo, S., Cafri, G., Rodriguez, J., and Rodriguez, M. 2006. "Relations among Multiple Peer Influences, Body Dissatisfaction, Eating Disturbance, and Self-Esteem: A Comparison of Average Weight, at Risk of Overweight, and Overweight Adolescent Girls," *Journal of Pediatric Psychology* (32:1), pp. 24-29 (doi: 10.1093/jpepsy/jsl022).

*Timmerman, J. G., Tonis, T. M., Marit, G. H. D.-v. W., Stuiver, M. M., Michel, W. J. M. W., van Harten, W. H., Hermens, H. J., and Vollenbroek-Hutten, M. M. R. 2016. "Co-Creation of an ICT-Supported Cancer Rehabilitation Application for Resected Lung Cancer Survivors: Design and Evaluation," *BMC Health Services Research* (16), p. 155 (doi: 10.1186/s12913-016-1385-7).

Titah, R., and Barki, H. 2009. "Nonlinearities between Attitude and Subjective Norms in Information Technology Acceptance: A Negative Synergy?," *MIS Quarterly* (33:4), pp. 827-844 (doi: 10.2307/20650329).

*Tregarthen, J. P., Lock, J., and Darcy, A. M. 2015. "Development of a Smartphone Application for Eating Disorder Self-Monitoring," *International Journal of Eating Disorders* (48:7), pp. 972-982 (doi: 10.1002/eat.22386).

*Tsai, C. C., Lee, G., Raab, F., Norman, G. J., Sohn, T., Griswold, W. G., and Patrick, K. 2007. "Usability and Feasibility of PmEB: A Mobile Phone Application for Monitoring Real Time Caloric Balance," *Mobile Networks and Applications* (12:2-3), pp. 173-184 (doi: 10.1007/s11036-007-0014-4).

*Tsanas, A., Saunders, K. E. A., Bilderbeck, A. C., Palmius, N., Osipov, M., Clifford, G. D., Goodwin, G. M., and De Vos, M. 2016. "Daily Longitudinal Self-Monitoring of Mood Variability in Bipolar Disorder and Borderline Personality Disorder," *Journal of Affective Disorders* (205), pp. 225-233 (doi: 10.1016/j.jad.2016.06.065).

*Tu, A. W., Watts, A. W., Chanoine, J.-P., Panagiotopoulos, C., Geller, J., Brant, R., Barr, S. I., and Mâsse, L. 2017. "Does Parental and Adolescent Participation in an E-Health Lifestyle Modification Intervention Improves Weight Outcomes?," *BMC Public Health* (17:1), p. 352 (doi: 10.1186/s12889-017-4220-0).

Turel, O. 2015. "Quitting the Use of a Habituated Hedonic Information System: A Theoretical Model and Empirical Examination of Facebook Users," *European Journal of Information Systems* (24:4), pp. 431-446 (doi: 10.1057/ejis.2014.19).

*Turk, M. W., Elci, O. U., Wang, J., Sereika, S. M., Ewing, L. J., Acharya, S. D., Glanz, K., and Burke, L. E. 2013. "Self-Monitoring as a Mediator of Weight Loss in the SMART Randomized Clinical Trial," *International Journal of Behavioral Medicine* (20:4), pp. 556-561 (doi: 10.1007/s12529-012-9259-9).

*Turner-McGrievy, G. M., Beets, M. W., Moore, J. B., Kaczynski, A. T., Barr-Anderson, D. J., and Tate, D. F. 2013. "Comparison of Traditional Versus

Mobile App Self-Monitoring of Physical Activity and Dietary Intake among Overweight Adults Participating in an mHealth Weight Loss Program," *Journal of the American Medical Informatics Association* (20:3), pp. 513-518 (doi: 10.1136/amiajnl-2012-001510).

*Turner-McGrievy, G. M., Wilcox, S., Boutte, A., Hutto, B. E., Singletary, C., Muth, E. R., and Hoover, A. W. 2017. "The Dietary Intervention to Enhance Tracking with Mobile Devices (Diet Mobile) Study: A 6-Month Randomized Weight Loss Trial," *Obesity* (25:8), pp. 1336-1342 (doi: 10.1002/oby.21889).

Tzoulaki, I., Elliott, P., Kontis, V., and Ezzati, M. 2016. "Worldwide Exposures to Cardiovascular Risk Factors and Associated Health Effects: Current Knowledge and Data Gaps," *Circulation* (133:23), pp. 2314-2333 (doi: 10.1161/CIRCULATIONAHA.115.008718).

*Umapathy, H., Bennell, K., Dickson, C., Dobson, F., Fransen, M., Jones, G., and Hunter, D. J. 2015. "The Web-Based Osteoarthritis Management Resource My Joint Pain Improves Quality of Care: A Quasi-Experimental Study," *Journal of Medical Internet Research* (17:7), p. e167.

*Vaughn-Cooke, M., Nembhard, H. B., Ulbrecht, J., and Gabbay, R. 2015. "Informing Patient Self-Management Technology Design Using a Patient Adherence Error Classification," *Engineering Management Journal* (27:3), pp. 124-130.

*Velardo, C., Shah, S. A., Gibson, O., Clifford, G., Heneghan, C., Rutter, H., Farmer, A., Tarassenko, L., and Team, E. C. 2017. "Digital Health System for Personalised Copd Long-Term Management," *BMC Medical Informatics & Decision Making* (17), pp. 1-13.

*Verdezoto, N., and Gronvall, E. 2016. "On Preventive Blood Pressure Self-Monitoring at Home," *Cognition Technology & Work* (18:2), pp. 267-285.

*Vogel, J., Auinger, A., Riedl, R., Kindermann, H., Helfert, M., and Ocenasek, H. 2017. "Digitally Enhanced Recovery: Investigating the Use of Digital Self-Tracking for Monitoring Leisure Time Physical Activity of Cardiovascular Disease (Cvd) Patients Undergoing Cardiac Rehabilitation," *PLoS ONE* (12:10).

Wagner, E. H., Austin, B. T., Davis, C., Hindmarsh, M., Schaefer, J., and Bonomi, A. 2001. "Improving Chronic Illness Care: Translating Evidence into Action," *Health affairs* (20:6), pp. 64-78.

Wagner, W., and Hayes, N. 2005. *Everyday Discourse and Common Sense: The Theory of Social Representations*. Palgrave Macmillan.

Walker, J. 2001. *Control and the Psychology of Health*. Open University Press, Buckingham.

*Wang, J., Sereika, S. M., Chasens, E. R., Ewing, L. J., Matthews, J. T., and Burke, L. E. 2012. "Effect of Adherence to Self-Monitoring of Diet and Physical Activity on Weight Loss in a Technology-Supported Behavioral Intervention," *Patient Preference and Adherence* (6), pp. 221-226.

*Webber, K. H., Tate, D. F., Ward, D. S., and Bowling, J. M. 2010. "Motivation and Its Relationship to Adherence to Self-Monitoring and Weight Loss in a 16-Week Internet Behavioral Weight Loss Intervention," *Journal of Nutrition Education and Behavior* (42:3), pp. 161-167.

Webster, J., and Watson, R. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Quarterly* (26:2), pp. xiii-xxiii.

Weingarten, S. R., Henning, J. M., Badamgarav, E., Knight, K., Hasselblad, V., Gano Jr, A., and Ofman, J. J. 2002. "Interventions Used in Disease Management Programmes for Patients with Chronic Illnesswhich Ones Work? Meta-Analysis of Published Reports," *BMJ* (325:7370), p. 925.

*Welch, J. L., Astroth, K. S., Perkins, S. M., Johnson, C. S., Connelly, K., Siek, K. A., Jones, J., and Scott, L. L. 2013. "Using a Mobile Application to Self-Monitor Diet and Fluid Intake among Adults Receiving Hemodialysis," *Research in Nursing & Health* (36:3), pp. 284-298.

*Welch, J., Dowell, S., and Johnson, C. S. 2007. "Feasibility of Using a Personal Digital Assistant to Self-Monitor Diet and Fluid Intake: A Pilot Study," *Nephrology Nursing Journal* (34:1), pp. 43-48.

West, P., Van Kleek, M., Giordano, R., Weal, M., and Shadbolt, N. 2017. "Information Quality Challenges of Patient-Generated Data in Clinical Practice," *Frontiers in Public Health* (5).

Westra, H. A., Dozois, D. J., and Marcus, M. 2007. "Expectancy, Homework Compliance, and Initial Change in Cognitive-Behavioral Therapy for Anxiety," *Journal of consulting and clinical psychology* (75:3), p. 363.

*Wharton, C. M., Johnston, C. S., Cunningham, B. K., and Sterner, D. 2014. "Dietary Self-Monitoring, but Not Dietary Quality, Improves with Use of Smartphone App Technology in an 8-Week Weight Loss Trial," *Journal of Nutrition Education and Behavior* (46:5), pp. 440-444.

WHO. 2014. "Global Status Report on Noncommunicable Diseases 2014." from http://www.who.int/nmh/publications/ncd-status-report-2014/en/

WHO. n.d. "Integrated Chronic Disease Prevention and Control." from http://www.who.int/chp/about/integrated_cd/en/

Wilde, M. H., and Garvin, S. 2007. "A Concept Analysis of Self-Monitoring," *Journal of Advanced Nursing* (57:3), pp. 339-350.

*Williamson, D., Anton, S., Han, H., Champagne, C., Allen, R., LeBlanc, E., Ryan, D., Rood, J., McManus, K., Laranjo, N., Carey, V., Loria, C., Bray, G., and Sacks, F. 2010. "Early Behavioral Adherence Predicts Short and Long-Term Weight Loss in the Pounds Lost Study," *Journal of Behavioral Medicine* (33:4).

Wilson, E. V., Mao, E., and Lankton, N. K. 2010. "The Distinct Roles of Prior IT Use and Habit Strength in Predicting Continued Sporadic Use of IT," *Communications of the Association for Information Systems* (27:12), pp. 185-206.

*Wolin, K. Y., Steinberg, D. M., Lane, I. B., Askew, S., Greaney, M. L., Colditz, G. A., and Bennett, G. G. 2015. "Engagement with Ehealth Self-Monitoring in a Primary Care-Based Weight Management Intervention," *PLoS ONE* (10:10), pp. 1-13.

*Yon, B. A., Johnson, R. K., Harvey-Berino, J., and Gold, B. C. 2006. "The Use of a Personal Digital Assistant for Dietary Self-Monitoring Does Not Improve the Validity of Self-Reports of Energy Intake," *Journal of the American Dietetic Association* (106:8), pp. 1256-1259.

*Young, L. A., Buse, J. B., Weaver, M. A., Vu, M. B., Mitchell, C. M., Blakeney, T., Grimm, K., Rees, J., Niblock, F., Donahue, K. E., and Monitor Trial, G. 2017. "Glucose Self-Monitoring in Non-Insulin-Treated Patients with Type 2 Diabetes in Primary Care Settings a Randomized Trial," *JAMA Internal Medicine* (177:7), pp. 920-929

Zhang, P. 2013. "The Affective Response Model: A Theoretical Framework of Affective Concepts and Their Relationships in the ICT Context," *MIS Quarterly* (37:1), pp. 247-274.

*Zhu, H. N., Luo, Y. H., Choe, E. K.. 2017. Making Space for the Quality Care: Opportunities for Technology in Cognitive Behavioral Therapy for Insomnia." *In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 5773-5786). ACM

Chapter 2 - Essay 2

Human Agents, Machine Agents and User Commitment on Digital Platforms: A Mechanism-Based Meta-Schema

Jinglu Jiang
Department of Information Technologies, HEC Montreal
3000, chemin Côte-Sainte-Catherine
Montréal, QC
H3T 2A7
CANADA
jinglu.jiang@hec.ca

Suzanne Rivard
Department of Information Technologies, HEC Montreal
3000, chemin Côte-Sainte-Catherine
Montréal, QC
H3T 2A7
CANADA
suzanne.rivard@hec.ca

Ann-Frances Cameron
Department of Information Technologies, HEC Montreal
3000, chemin Côte-Sainte-Catherine
Montréal, QC
H3T 2A7
CANADA
ann-frances.cameron@hec.ca

Abstract

As digital platforms have become a widespread way to transact, the relationships between representations of platform offerings, user commitment, and collective outcomes have attracted researchers' attention. Notwithstanding the contributions of extant research in identifying the antecedents of user decisions on digital platforms, we seek to address two issues that may hinder the accumulation of knowledge and further development of the field. First, the research is limited to human agents as actors and has ignored the role of machine agents, despite their increasing presence on platforms. Second, the majority of the research adopts a variance-based theorizing approach, which does not reveal the cogs and wheels operating among the related constructs. Accordingly, we adopt a mechanism-based approach to develop a meta-schema of seven mechanisms that come into play to explain how the representation of platform offerings affects agent actions (both human agents and machine agents) and how those agent actions at lower levels give rise to collective outcomes. By simultaneously considering social mechanisms to explain actions by human agents and computational mechanisms to explain actions by machine agents, our meta-schema provides a vocabulary and constitutes a canvas that can be leveraged by future research.

Keywords: digital platform, user commitment, algorithm, machine agents, mechanism, theory-building

2.1 Introduction

Contemporary digital platforms serve as a new way of organizing in which a swirl of information streams converge to form seemingly transparent signals of both the qualities of the offerings— the products, services, or funding campaigns that are offered on digital platforms—and the social dynamics that surround them. Indeed, many digital platforms connect to social media (e.g., links to share offering details on Facebook and Twitter), allow crowd evaluation of the offering (e.g., commenting and rating), and display ongoing updates about the offering (e.g., total sales, total funds raised, highest bid, recent interactions with the offering providers, and relevant news). These in turn influence user perceptions of these offerings, their behaviors on digital platforms, and collective outcomes that emerge (Thies et al. 2016; Levina and Arriaga 2014; Orlikowski and Scott 2015; Viglia et al. 2018). Our research seeks to examine this phenomenon, specifically, the relationships between representations of platform offerings, user commitment, and collective outcomes.

This phenomenon has attracted a significant amount of research, which has advanced our understanding of the extent to which user perceptions and behaviors are influenced by antecedents such as offering characteristics (e.g., the impact of crowdfunding campaign characteristics on an individual's funding decisions, Hong et al. 2018), user-generated content (e.g. the impacts of peer reviews on consumers' purchasing decisions, Wang et al. 2018), social influences (e.g. herding effect and eWoM in online sales, Li and Wu 2018), and website features (e.g., offering recommendation style, Benlian et al. 2012).

Notwithstanding the contributions of extant research, we endeavor to expand current knowledge by addressing two issues that may hinder further development of the field. First, our theorizing accounts for a critical characteristic of contemporary platforms, the presence of algorithms—in particular of machine agents—that can use the data generated by human agents to modify the offering representations, thus playing an important role in influencing agent actions and collective outcomes. Indeed, extant research largely assumes that data presented on digital platforms (e.g., product reviews, ratings, transaction records) are generated by human agents, thus reflecting human opinions and behaviors. Recently, however, there has been a significant increase in the presence of machine agents (e.g., chatterbots, automatic transaction algorithms, virtual assistants, recommendation agents) that can also generate content on digital platforms. For instance, bot traffic accounts for over 40% of all web traffic (DistilNetworks 2018) and Facebook disabled almost 13 billion fake accounts in 2018, many of them bots (Recode 2018). Moreover, new technological advances such as machine learning and artificial intelligence (AI) also play an increasingly important role that has profoundly influenced the relationships between humans, machines, and platforms (Rai et al. 2019). It is reported that 61% of businesses implemented some form of AI in 2017, including predictive analytics, natural language processing, voice recognition and recommendation engines (NarrativeScience 2018). The chatterbots are learning user behaviors in an attempt to interact with human users as if they are also human (e.g., Microsoft's AI chatterbot "Xiaoice"). Virtual personal assistants are applying advanced voice recognition and natural language processing techniques to fully automate previously manual tasks (e.g., Taco Bell and Domino's food ordering bot, Sephora's makeup suggestion assistant). These advanced machine agents use various sources of data (including real-time user-generated data) to perform tasks, sometimes generating information that is not distinguishable from human-generated content. It is thus essential that theorizing on the impact of offering representations on digital platforms account for the presence and the role of both machine and human agents.

Second, we adopt a mechanism-based approach, comprising social mechanisms and computational mechanisms, to explain our phenomenon. In doing this, we address the second issue of extant research, which usually adopts a variance approach to explanation that emphasizes the role of the variations in antecedents to explain the variations in an outcome. This approach has undeniably contributed to our understanding of key antecedents of human agent actions and collective outcomes. We contend, however, that the complex and dynamic nature of digital platforms calls for a theorizing approach aimed at explaining why outcomes at the platform level occur by describing how human and machine agents' actions result in collective outcomes. Espousing the view that new theorization on human-machine hybrids on digital platforms is needed (Rai et al. 2019), we go beyond a modularized understanding of human cognition, behaviors, and platform infrastructures, and mobilize a mechanism-based approach as a foundation for developing our explanation.

Our objective is to theorize the mechanisms by which digital platform offerings—through the actions of human and machine agents—influence user commitment and collective outcomes. We focus on two types of collective outcomes (i.e., commitment and non-committal outcomes) which emerge from individual-level actions. Collective commitment outcomes result from commitment actions (i.e., a binding obligation that can take the form of a promise, payment, or bid related to the offerings, such as purchasing a product, making a reservation, or contributing to an online campaign). Non-committal outcomes result from non-committal actions, which are non-binding actions that can take the form of content creation, relationship building, or evaluation of the offering. As multi-sided digital platforms have become an increasingly popular way to transact (Constantinides et al. 2018), we develop our explanation in the particular context of multi-sided digital platforms where money-related user commitment is the primary objective.

To fulfill our objective, we adopt a mechanism-based approach to explanation (Hedström and Swedberg 1998; Piccinini 2007; Illari and Williamson 2012). We theorize the causal capacities of the offerings, the causal mechanisms that generate agent actions (i.e., human agents and machine agents), the causal mechanisms that update the causal capacities through algorithms, and the associated outcomes. Drawing on social mechanisms to understand actions by human agents and on computational mechanisms to understand actions by machine agents, we develop a meta-schema[6] that comprises seven key mechanisms: human agent cognitive frame formation, machine agent belief formation, human agent action formation, machine agent action selection, causal capacity update, offering representation update, and collective outcome emergence.

Our study makes several contributions. First, our work responds to calls for research that recognizes and explores the roles of both humans and machines on digital platforms (Rai

[6] A schema is "a truncated abstract description of a mechanism that can be filled with more specific descriptions of component entities and activities" (Darden 2002 p. S356). A meta-schema is a high-level schema from which other schemas can be derived.

et al. 2019). In addition to examining human agents, it actively acknowledges and seeks to explore the influence of both machine agents and platform-based algorithms on collective outcomes. Second, our study goes beyond the variance-based approach currently used in this research area, instead employing a mechanism-based approach to explain the cogs and wheels that bring about relationships between offering, agent actions, and outcomes (Avgerou 2013). Finally, our meta-schema serves as a platform for future research as it provides a common vocabulary of important constructs, delineates seven important mechanisms related to human and machine agent behaviors, and serves as a foundation on which researchers can build more contextualized theories. We briefly illustrate how researchers can leverage our meta-schema by proposing three different approaches – difference-making, agent modeling, and data-driven discovery – in the hope of inspiring researchers to go beyond human agents and to actively incorporate the potentially powerful influence of machine agents and platform algorithms in their future work.

2.2 Existing Literature: Impact of Digital Platform Offering Representation on User Commitment and Outcomes

In order to understand the type of research that is currently being conducted on money-related user commitment in the context of multi-sided digital platforms, we reviewed studies published in journals on the AIS Senior Scholar Basket of Eight list. We sought studies based on three criteria. First, the studies examine multisided platforms where user-generated content about the offerings is visible to others (e.g., reviews, ratings, past purchases). Second, monetary user commitment should be empirically examined either as an independent, intermediary or dependent construct. In line with our provided definitions, user commitment had to take the form of a binding obligation toward an offering, and the offering could be a product or service. Third, user commitment[7] should be conceptualized and measured at either the individual level or at the collective level.

[7] Consumer commitment is a related term used in e-commerce research. However, it is a complex construct with an affective aspect such as the degree of liking, a cognitive aspect such as preference and a satisfaction and behavioral aspect such as repetitive purchasing and maximum effort to maintain the relationship (Bloemer and Kasper 1995; Bilgihan and Bujisic 2015; Eastlick et al. 2006; Oliver 1999).

Thirty-nine studies were identified, with the vast majority of them being published within the last ten years. The study profile that shows publication distribution and the platform investigated is presented in Appendix B (see Table B1 and B2). Online retail is the most frequently studied context (N=14, e.g., Carmi et al. 2017; Guo et al. 2018; Lin et al. 2017). However, recent years have witnessed an increasing interest in online auctions (N= 6, e.g., Hinz et al. 2016, Reiner et al. 2014), P2P lending (N=4, e.g., Feller et al. 2017; Jiang et al. 2018; Xu and Chau 2018), and crowdfunding (N=8, e.g., Burtch et al. 2016; Thies et al. 2018; Zheng et al. 2018). Other platforms examined include group-purchasing, and online knowledge markets. The literature also exhibits a wide variety in terms of the type of offering (e.g., crowdfunding campaign, online consultation service, goods and products), the business model that delivers the offering (e.g., B2C model, C2C model, professional service delivery), and types of commitment activities (e.g., purchase, repurchase, campaign contribution).

Offering representation (i.e., the description and presentation of the offering) is one of the most studied antecedents (see Appendix B Table B3). We conceptualize this term broadly, not limiting it to the offering itself but also to include any elements surrounding it, such as the information related to the provider and brand. These elements are integral to the offering and contribute to the users' evaluation of the offering. The most frequently examined offering representations on digital platforms include product or service characteristics such as price (e.g., Hu et al. 2017), offering reviews (e.g., Wu et al. 2013), social media and electronic word of mouth (eWOM) related to the offering (Gu et al. 2012), offering provider characteristics such as social capitals and country of origin (e.g., Hong and Pavlou 2017; Hong et al. 2018), provider activities such as feedback and interacting with customers (e.g., Ou et al. 2014; Zheng et al. 2018), peer activities related to the offering (e.g., Burtch et al. 2016) and platform-implemented features that influence presentation of all offerings (e.g., Kuan et al. 2014; Reiner et al. 2014; Thies et al. 2018). The majority of the studies examined direct impacts of offering representations on individual commitment or collective outcomes. Although some studies include

Although user cognition is an important source to drive actions, the current study focuses on the behavioral aspect of commitment.

explanations related to the intermediate mechanisms linking the different components (such as how offering representation influences user beliefs and attitudes, which in turn influences user action), only a few empirically incorporate these intermediate mechanisms in their analytical model (Ou et al. 2014 is an notable exception).

Various user beliefs, attitudes and perceptions are also frequently examined as antecedents of commitment actions. Among them utility-related antecedents (e.g., perceived ease of use, perceived effectiveness, satisfaction), trust and risk perceptions are most widely tested (e.g., Benlian et al. 2012; Huang et al. 2017; Ou et al. 2014; Zheng et al. 2018). Besides the direct impact of user perceptions on behaviors, some studies investigated the chain of perception which provide opportunities to uncover finer-grained mechanisms that give rise to users' commitment actions. For example, Pavlou and Gefen (2005) examined how buyer's psychological contract violations may influence perceptions of risk and trust. A summary of the key antecedent categories for user commitment is presented in Table 2.1 below.

Table 2.1. Key Antecedents of User Commitment

Antecedent Category	Example Study
Offering Representation	
Product/ service characteristics	Burtch et al. (2013), Burtch et al. (2016), Burtch et al. (2018), Chen et al. (2015), Ghose et al. (2006), Gleasure et al. (2017), Hong et al. (2018), Hu et al. (2017), Jiang et al. (2018), Li and Wu (2018), Lin et al. (2017), Liu et al. (2014), Özpolat et al. (2013), Pavlou and Gefen (2005), Wu et al. (2013)
Offering provider characteristics	Feller et al. (2017), Gregg and Walczak (2008), Guo et al. (2017), Hong and Pavlou (2017), Kim and Ahn (2007)
Provider activities	Benlian et al. (2012), Gleasure et al. (2017), Guo et al. (2017), Hong et al. (2018), Ou et al. (2014), Xu and Chau (2018), Zheng et al. (2018)
Peer activities	Forman et al. (2008), Burtch et al. (2013), Ge et al. (2017), Gu et al. (2012), Hu et al. (2017), Huang et al. (2017), Kuan et al. (2014), Li and Hitt (2008), Li and Wu (2018), Lin et al. (2017), Liu et al. (2014), Thies et al. (2016), Wu et al. (2013), Xu and Chau (2018)
Platform-implemented features	Burtch et al. (2018), Jiang et al. (2018), Kim and Ahn (2007), Oestreicher-Singer and Sundararajan (2012), Ou et al. (2014), Özpolat et al. (2013), Reiner et al. (2014), Thies et al. (2018)
Other	Hinz et al. (2016)
User Beliefs, Attitudes, Perceptions	

Utility-related	Guo et al. (2018), Huang et al. (2017), Kuan et al. (2014), Pavlou and Gefen (2004), Pavlou and Gefen (2005), Xu and Chau (2018), Zheng et al. (2018)
Trust/ Risk-related	Guo et al. (2018), Kim and Ahn (2007), Ou et al. (2014), Pavlou and Gefen (2004), Pavlou and Gefen (2005), Van Slyke et al. (2006), Wu et al. (2013)

Individual-level actions are diverse. User commitment actions include crowdfunding campaign contributions (e.g., Burtch et al. 2013), repurchases (e.g., Ou et al. 2014), loan payments (e.g., Ge et al. 2017), service contract selections (e.g., Hong and Pavlou 2017) and general transactions (e.g., Kim and Ahn 2007). Non-committal actions include reviews (Benlian et al. 2012), likes (Quan et al. 2014), and eWOM (Gu et al. 2012). In addition, a variety of collective outcomes were examined including total contribution to a campaign (e.g., Hong et al. 2018), total sales (e.g., Lin et al. 2017), total revenue (e.g., Oestreicher-Singer and Sundararajan 2012), and average reviewer rating (Hu et al. 2017).

The studies predominantly adopt a variance-based approach such as econometric modeling and structural equation modeling to explain the impacts of offering representations (one notable exception is Gleasure et al. (2017) who used a sociomaterial lens to examine a case study). Although the majority of the studies do provide explanations during hypothesis development or in the interpretation of results, what happens between the inputs (or independent variables) and outputs (or dependent variable) is often not directly examined and, thus remains a black box.

2.3 Conceptual Background: Mechanism-Based Explanation

As shown in the literature review, extant studies mainly adopt a variance-based approach to explanation. This approach has undeniably contributed to our understanding of user commitment on platforms by identifying the key antecedents of commitment. We contend, however, that the complex and dynamic nature of digital platforms calls for a theorizing approach aimed at explaining *why* commitment actions and collective outcomes occur by describing *how* platform components and their associated agents—be they human or machine—act and interact. We suggest a mechanism-based approach as an appropriate foundation for developing this type of explanation. As such, a mechanism-based explanation can complement the extant variance-based approach as it supports the theorization of the causal processes that link the explanandum and the statistically relevant explanans of a variance-based theory (Avgerou 2013, Salmon 1984a). It should be noted that the mechanism-based explanation overlaps with process-based and activity-based explanations, and the actual testing of the resulting theory can be done using quantitative or qualitative approaches.

Current conceptualizations of mechanisms mainly originate from two fields. First, several philosophers of science (e.g., Bechtel 2005, 2009; Craver 2007; Glennan 2011; Machamer et al. 2000) promoted the New Mechanism paradigm that has been mostly applied in life sciences such as biology, cognitive science, and neuroscience. In parallel, sociologists (e.g., Elster 1998; Hedström and Bearman 2009; Little 2011) developed a social mechanism paradigm, which is widely used in sociology, political science and economics (Psillos 2011). Although the two conceptualizations originate from different disciplines, they share several constitutive elements (Illari and Glennan 2017). Indeed, both conceptualizations view mechanisms as complex systems or processes that have constitutive and causal dimensions, are embedded in context, and reflect reality. Because our theorizing effort pertains to a hybrid system that involves both human and machine agents, we deem that adopting an integrative view of mechanisms should help to develop a rich explanation of the interactions of human agents and machine agents through digital platforms.

2.3.1 The Nature of Mechanisms

Although the literature proposes many definitions of mechanism, most of them comprise the following metaphysical elements: a phenomenon, entities, activities and their organization, and the function of the mechanism (Craver and Tabery 2017). In this work, we adopt Illari and Williamson's (2012) definition: "a mechanism for a phenomenon consists of entities and activities organized in such a way that they are responsible for the phenomenon" (p. 120).

A mechanism pertains to a particular phenomenon, which sets the boundary of the mechanism: what is in the mechanism and what is not should always refer to the phenomenon to be explained, and the entities and their activities in a mechanism should be related to the phenomenon (Illari and Williamson 2012). Mechanisms are responsible for the phenomenon in three ways (Craver and Darden 2013). A mechanism can *produce* the phenomenon in that an object, a state or an event is produced by some causal sequences (e.g., the photosynthesis mechanism in which plants turn sunlight, carbon dioxide and water into energy). A mechanism can *underlie* the phenomenon, in that the parts of a mechanism are organized in certain ways to give rise to the phenomenon (e.g., the neuron activities that underlie the working memory constitutes mechanisms for maintaining information in the brain). Lastly, a mechanism can *maintain* the phenomenon, ensuring a state of equilibrium, correcting any deviation from the equilibrium point (e.g., the regulation mechanism of body temperature).

Second, mechanisms are composed of entities—i.e., components or parts—and their activities (Glennan 1996). Entities are the producers of changes. Entities are identified by their properties (e.g., locations, structures and orientations) that enable their engagement in the activities (e.g., push, pull, transmit, give feedback). Although some researchers suggest that entities should be stable (Glennan 2002), others view entities and their properties as unstable and emergent (Illari and Williamson 2012). In short, searching for a mechanism is, in part, identifying the working entities and mapping entities to activities (Bechtel and Abrahamsen 2005).

Third, a mechanism is organized such that its entities and activities are set up to do something (Craver 2001). The entities have a spatiotemporal organization (e.g., location, shape, position, order, duration, rate, interaction direction) and are working with startup, ongoing, and termination conditions[8]. The organization can take various forms, such as linear (i.e., the completion of the first stage gives rise to the second), in cycles (i.e., key products exist at juncture stages and can be reused cyclically), in networks (i.e., clusters of units are causally connected), additive (i.e., the whole changes linearly with the addition of the parts), or emergent (i.e., the whole is more than the sum of the parts, and new properties may appear). The same entities and activities may produce different

[8] Not all mechanisms have a termination condition, and a single input or output without a final termination state is possible.

outcomes if organized differently (Craver and Tabery 2017, Levy and Bechtel 2013, Wimsatt 1997).

Finally, the entities are organized so as to carry out activities that accomplish the *function* of the mechanism (Cummins 1975, Garson 2013). For example, the heart is beating to circulate the blood, the car engine burns fuel to produce mechanical power, and the web crawler for search engines automatically fetches webpage scripts for indexing. By specifying the functional characterization of a mechanism, we can reveal why a mechanism is there (Bechtel and Abrahamsen 2005).

2.3.2 Mechanism-Based Explanation

Mechanism-based explanation refers to explaining *why* a phenomenon happens by describing *how* some mechanisms produce the phenomenon (Halina 2017). It identifies the entities and activities that are responsible for a phenomenon and determines their organization. Thus, mechanism-based explanations identify the working entities that constitute the phenomenon and show how the parts work to cause the phenomenon. Accordingly, mechanism-based explanation involves both constitutive explanation (i.e., showing causal structure by clarifying the causal properties of the parts and their organization) and etiological explanations (i.e., showing causal processes by showing the chains of events and their antecedent causes, such as activity-based explanation), which delimits relevant components and causal processes responsible for the phenomenon (Craver 2007, Kuorikoski 2012, Salmon 1984b, Ylikoski 2013).

The issue of whether a mechanism-based explanation should be ontic or epistemic was debated in the literature. Some authors held an ontic view (e.g., Craver 2001) and matched a phenomenon into causal structures of the world. Thus the explanation *exhibits* the objective portions of the causal structure (i.e., the entities and activities) that bring about the phenomenon. An ontic explanation is thus a genuine causal structure of the real world (e.g., the entities and activities that form the mechanism). Others held an epistemic view (e.g., Bechtel and Abrahamsen 2005), whereby a mechanism-based explanation is the understanding of a mechanism by scientists, so it is the communicative acts that convey information about the mechanism of a phenomenon that is explanatory—that is, the mechanism itself is not a mechanism-based explanation, the description of the mechanism is. In line with more recent discussions (Illari 2013), we hold an integrative view of this ontic-epistemic distinction, whereby the explanation should reflect the nature of reality, but the process of explaining involves idealization and abstraction which constrain and are constrained by researchers' goals and communicative acts (Halina 2017, Wright 2012).

In IS, a few studies use generative mechanisms (Bygstad et al. 2016, Henfridsson and Bygstad 2013) to explain digital infrastructure evolution. Despite the epistemological assumption (i.e., the strong root in critical realism) and the difference in study object (i.e., digital platform generativity), the key components and characteristics of a generative mechanism are similar to those of a mechanism we discuss here – the causal structures including entities and activities, and the processes that generate the outcomes (Henfridsson and Bygstad 2013). Whereas future context-specific researches can choose their own ontological and epistemological stances, our current study applies a general conceptualization of mechanism-based explanation that emphasizes common components and functions across areas of science.

2.3.3 Multilevel Nature of Mechanisms and Interlevel Causation

In terms of organization, mechanisms are multilevel, because the behavior of the mechanism as a whole (i.e., the higher level) can be decomposed into the behaviors of its components (i.e., the lower level). For example, the behavior of a society can be decomposed into the behaviors of collectives, which can be decomposed into the behavior of individuals, which can be further decomposed into the behaviors of individual organisms, which can then be decomposed into the physiological functions of the organisms, and so on. The lower-level properties are realizers of higher-level functions so that the parts (i.e., the lower-level entities) interact and give rise to the behavior of the whole (Glennan 2010b). The levels of the mechanism are different levels of description for the same object. For instance, we can describe the same cognitive model from higher-order computational rules, the concrete input-output transformation algorithms, and the physical implementation in the brain (Zednik 2017). Such a decomposing process is guided by research objectives and is stopped when the decompositions are no longer relevant or limited by scientific progression (e.g., physiological mechanisms may not be relevant for behavioral researchers). Similarly, higher-level mechanisms may not be of interest to everyone. For instance, researchers on individual IS use researchers may not be interested in how the implementation of IT causes a social movement, whereas a sociologist may have great passion regarding the societal level impacts.

The multi-level nature of mechanisms highlights the importance of understanding how the parts interact with their whole, which raises the challenge of understanding interlevel causation (Craver and Bechtel 2007). The mechanism-based approach of making causal claims suggests that mechanisms are mediating causal interactions, and both bottom-up and top-down causation can occur in the hierarchy of mechanisms.

2.3.4 Social Mechanisms

Social mechanisms are defined as "social processes having designated consequences for designated parts of the social structure" (Merton 1968, cited in Hedström and Swedberg 1998 p.6). They explain how social-level causes generate social-level outcomes through social processes. Social causes involve the acts of individuals within social constraints, such as the rules of social institutions and actions of other individuals (Little 2011). The entities of a social mechanism include individuals, collectives, artifacts, and hybrids such as firms. Entities engage in sequences of activities that unfold over time to generate observed outcomes in their designated social contexts (Avgerou 2013).

Figure 2.1 depicts a typology of social mechanisms that Hedström and Swedberg (1998) developed according to Coleman's (1986) boat model. Although not using the term mechanism, Coleman's boat model indeed applies mechanism-based thinking (Ylikoski 2017). The boat model posits that to explain the macro-social reality (e.g., unemployment), simply identifying the associated macro-level social variables (e.g., the implementation of social welfare program) is not enough to establish the causal link. In fact, it is the micro-level social structures (i.e., the activities and interactions of individuals) that bring about the macro-level reality. Hence, the boat model is a macro-micro-macro process of how social institutions influence how individuals think and behave, and individuals' thinking and behaviors provide an explanatory understanding of macro-level social formation (Coleman 1986). Based on this conceptualization, Hedström and Swedberg (1998) posit that to explain a macro-level social phenomenon (i.e., arrow 4 on Figure 2.1), the macro level factors must be connected to the micro-level activities by individual actors. Three types of mechanisms come into play:

- *Situational mechanisms (arrow 1)*: social context influences individual actors' beliefs, habits and cognitive frames that constrain and enable these actors' opportunities for action.
- *Action-formation mechanisms (arrow 2)*: individual actors' opportunities and cognitive frames lead to or change these actors' behaviors.
- *Transformational mechanisms (arrow 3)*: individual actors' actions produce social patterns and outcomes (e.g., the emergent macro-level properties).

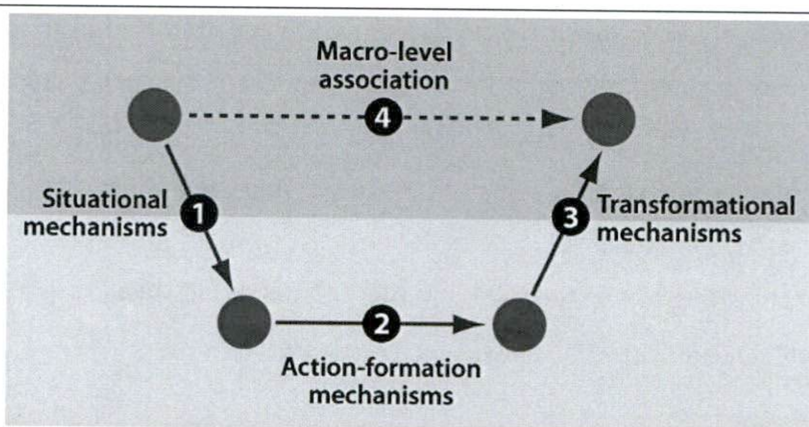


Figure 2.1 A Typology of Social Mechanisms Reprinted from Hedström and Ylikoski (2010) p.59

These three types of social mechanisms highlight the top-down influence of social context, as well as how social outcomes bottom out at the micro-level of agent actions. Different from a variance-approach to explanation that makes direct associations between social context and social outcomes (arrow 4), it is the underlying role of agents' beliefs and opportunities for actions and changes in behaviors that bring about the macro level outcomes, not the social structure itself (Ylikoski 2017). Depending on the research goal, researchers may not be interested in all three types of social mechanisms. For example, sociologists may not be interested in how individual cognitions transform into actions, whereas psychology or cognitive researchers may be highly interested in this mechanism.

2.3.5 Computational Mechanisms

Conceptualized within the New Mechanism paradigm, computational mechanisms are a particular type of mechanisms widely used to explain how physical and digital computational systems work. Neurophysiologists who work in computational neuroscience consider the brain as a computational system, and view the work performed by neurons as computation (Shagrir 2006). Cognitive scientists view humans and the workings of human mind as computation, that is, a process that manipulates mental representations (Putnam 1991).

Although computational mechanisms originally focused more on computation in the physical world, they have been mobilized to explain the operation and impacts of digital computers (and their algorithms). Indeed, computational mechanisms are deemed relevant in this context since digital computers are instructed (e.g., following the pre-arranged algorithms) to perform the desired task. Computer scientists design computer algorithms, which are a sequence of instructions that allow computers to solve problems within finite steps, and this problem-solving process is computation (Cormen et al. 2009).

Computational mechanisms are responsible for the functions and behaviors of computing systems—be they computers or human brains—by describing how "the input and output information streams are causally linked … along with the specific structure of information processing" (Miłkowski 2014, p.221). Together with the component parts and information processing structure, these input and output streams constitute the capacity of the mechanism (Piccinini 2010).

Computation can be viewed as a functional mechanism, since the entities of the computation mechanism are functionally organized to generate the capacity of the mechanism as a whole. Entities within a computational mechanism operate to deal with the designated functions. Computational mechanisms are real (i.e., the ontic aspect of a mechanism) and local to the empirical question that the algorithm is trying to solve (Illari and Williamson 2011).

2.4 A Mechanism Meta-Schema of Human Agents, Machine Agents and User Commitment on Digital Platforms

We now present our mechanism-based meta-schema of user commitment on digital platforms, which we developed drawing on Craver and Darden's (2013) mechanism discovery strategies. Because the mechanism discovery process begins with background knowledge about the phenomenon—here, the impact of digital platform offering representations on user commitment and collective outcomes—and the kinds of entities and activities that may be involved within the boundary of the phenomenon of interest, we developed a deep understanding of digital platforms. This was done by visiting, analyzing, using, and studying digital platforms. We specified the phenomenon by setting our contextual and conceptual assumptions (Rivard 2014), and we identified the key entities and activities, and their organization (Darden 1991, 2002).

Three types of mechanism schemas exist (Craver and Darden 2013): *how-possibly* schemas describe how a mechanism might work, *how-plausibly* schemas describe how a mechanism might work considering the known evidence, and *how-actually* schemas describe how the mechanism works in reality. Because of the theory-building nature of our work, we applied techniques such as observation, enumeration, and literature synthesis to construct the initial *how-possibly* and *how-plausibly* schema (Jaccard and Jacoby 2009). We also used modular subassembly (reasoning about how mechanism components might be combined) and forward-backward chaining (making inferences about what comes before and after a mechanism component) (Craver and Darden 2013) as thought experiments during the development process.

2.4.1 Underlying Assumptions

Our explanation is bounded by the following assumptions. First, we focus on multi-sided digital platforms that create value by coordinating the demands of two or more distinct groups of human agents—customers— "who need each other in some way" (Evans 2003 p.191). Second, while we do not limit the type of offerings being exchanged (i.e, we include both goods and services), we limit our theorization to platforms where monetary commitment is sought. Although other types of platforms that focus on non-monetary objectives share common elements with our context and could benefit from some part of our work (e.g., knowledge contribution to an online Q&A community), these other types of commitment do not exhibit the zero-sum nature of monetary commitment and our explanations may be inadequate in these contexts. Third, our explanation pertains to platforms where some of the human agents' actions—either their commitment (e.g., charitable donations, purchases of a good) or non-committal actions (e.g., reviews, shares, likes, comments)—leave a trace on the platform. Indeed, digital platforms are not showcases of static webpages that display offering information; they are virtual spaces with complex social interactions. Thus, we view social influence as an essential component to produce user commitment or non-committal actions, which in turn influence collective outcomes. Such social influence should be perceptible in some way as a signal. Although the exposure of peer actions is important, however, not all streams of human agent activities on digital platforms are necessarily visible. The platforms can have various interface designs, and not all agent actions will be presented on the interface. For example, the platform can choose to present aggregative information (e.g., total sales and sales ranks), or embed information about a human agent action in the HTML template so that a machine-agent such as a wrapper (i.e., an algorithm that can scrap HTML script and extract its content) can make sense of it but human agents cannot.

Because digital platforms coordinate the demands of distinct groups of human agents, they are a collective-level structure to which social mechanism-based explanations are relevant. Therefore, as we adopt Hedström and Swedberg's (1998) typology of social mechanisms to theorize human agents on platforms, we espouse the authors' cognitivist assumption. Under this assumption, the representation of platform offerings provides some signals to induce human agents' cognitive frames, which in turn produce rational choices and actions.

2.4.2 Concepts of the Meta-Schema

We first introduce the concepts at the higher-level—the offering representation on digital platforms and the associated collective outcomes. Then we decompose the underlying entities at the lower-level—human agents and machine agents—and their activities before we introduce their organization. Table 2.2 lists the definitions of these essential concepts.

Table 2.2 Definitions of Key Concepts

Concept	Definition
Offering Representation	The description of the set of attributes of a given offering on the digital platform.
Actual causal capacity of the offering representation	The extent to which the quality of the offering is attractive or persuasive.
Actual stable structure	The actual attributes of the elements of an offering, which are not modifiable by agent actions.
Actual dynamic structure	The actual attributes of the elements of an offering, which are subject to modification by agent actions. It includes collective commitment and other aggregate outcomes of previous actions.
Displayed causal capacity of the offering representation	The presentation of the offering-related information on the platform.
Displayed stable structure	The presentation of the non-modifiable offering-related information on the platform.
Displayed dynamic structure	The presentation of the modifiable offering-related information on the platform.
Human agent	An individual who performs actions on digital platforms, be they the target audience or party at the origin of the offering, e.g., customer, supplier, service provider, customer service representative
General-purpose algorithm	A set of step-by-step computational instructions to achieve the desired task, such as sorting, ranking, classification and clustering, in a finite number of moves without the ability to publish new content on the platform.
Internal general-purpose algorithms	A general-purpose algorithm implemented by the platform owners.
External general-purpose algorithm	A general-purpose algorithm introduced by external third parties.
Machine agent	"An agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives" (Wooldridge 2009 p.5). We define machine agents as those agents that can publish new content on the platform and act like human agents.
External machine agent	A third-party introduced machine agent (i.e., external to the platform) that performs human-like actions on the focal platform.
Internal machine agent	A platform-designed machine agent (i.e., internal to the platform) that performs human-like actions on the focal platform.

Table 2.2 Definitions of Key Concepts	
Concept	**Definition**
Human agent cognitive frames	An individual's state of mind linking an object to its attributes, which represents the information he or she has about the object.
Machine agent beliefs	The knowledge built up by the machine agents about the environment (e.g., the significant changes on the platform).
Actions	Operations performed by human agents or machine agents on the digital platform to create content, build relationships, evaluate the offerings, and/or commit to the offering.
Commitment	One type of action which is a binding obligation that can take the form of a promise, payment or bid related to the offerings.
Non-committal actions	A non-binding action that can take the form of content creation, relationship building, or evaluation of the offering.
Collective outcome of agent actions	The coalescence of lower-level (either commitment or non-committal) outcomes.

The offerings, either goods or services, are the core of interaction and value creation on digital platforms. The representations of the offerings (i.e., the set of attributes of a given offering) serve as signals that are perceptible by human and machine agents on the platform (e.g., offering descriptions, popularity, and offering provider identity), which influence agents' decision-making and actions. We define an offering representation's qualification such as attractiveness and persuasiveness as its *causal capacity,* which is underlain by causal structures, defined as the attributes of the elements of an offering. The elements of an offering can include part of the offering itself (e.g., the product description) or anything related to the offering (e.g., offering provider's information, product network, past transaction history). The attributes of these elements are causally relevant to agents' cognitive processes to evaluate the offering and action possibilities.

We differentiate four types of causal structures (see Table 2.3): actual versus displayed, and stable versus dynamic (i.e., 2 by 2 types). Some attributes of an offering are not modifiable (i.e., stable) by human agents or machine agents during the interaction and transaction process (e.g., offering description), whereas others are dynamic, being subject to modification by agent actions (e.g., real-time reviews and ratings). Dynamic structures not only reflect the outcome of individual agent actions but also reflect collective outcomes of the human agent's and machine agents' previous actions. Herein, we define the *collective outcome of agent actions* as the coalescence of lower-level outcomes, including both commitment and non-committal outcomes. It is collective, thus going beyond individual-level results. It may have two statuses: the final outcome (e.g., the success of a crowdfunding campaign), or the temporary milestones that can be reached repeatedly over time (e.g., top 10 sales ranking). The latter contributes to an offering's dynamic structure, whereas the former is the end-state of a mechanism. Since not all offerings require a final outcome (e.g., a product that is continuously for sale), the mechanism may not have an end-state.

Table 2.3 Types of Offering Causal Structures

	Stable	Dynamic
Actual	Actual stable structure	Actual dynamic structure
Displayed	Displayed stable structure	Displayed dynamic structure

Human agents are individuals who perform actions on digital platforms, be they the target audience or party at the origin of the offering— including service or product providers, customers who pay for the services and products, and platform moderators who interact with providers and customers. However, we exclude platform administrative or technical personnel who are responsible for designing (and changing) the platform's architecture[9].

The platform ecosystem is supported by algorithms, defined as a set of step-by-step computational instructions to achieve the desired task. Algorithms can deal with various tasks with a wide range of complexity (e.g., simple information display, information screening and processing, dialogue with other algorithms and human users). Accordingly, we classify the algorithms into four types (see Table 2.4): those algorithms that have the ability to publish new information are defined as *machine agents*, and those that just have the ability to process and manipulate existing information are defined as *general-purpose algorithms*. According to intelligent agent literature, "an agent is a computer system that is situated in some environment, and that is capable of autonomous action in this

[9] We use this exclusion criterion because the type of actions they are performing (i.e., designing and modifying the platform technological design) is not the focus of the current study.

environment in order to meet its design objectives" (Wooldridge 2009 p. 5). Machine agents in some ways act like human agents. They are autonomous, so they can make independent decisions and operate without direct human intervention. They are rational in that they are goal-oriented and will not act in such a way as to prevent the goal from being achieved (Wooldridge and Jennings 1995). For an agent to be intelligent, it needs to be reactive, proactive and social (Wooldridge 2009). Reactivity is the ability of an intelligent agent to respond to changes in the environment; proactiveness is the ability of an intelligent agent to pursue goal-directed behavior by taking the initiative (i.e., persistence and robustness); and social ability is the performativity of an intelligent agent to interact with other agents and possibly humans (Wooldridge and Jennings 1995). Some researchers take a strong agency perspective and view intelligent agents as having mental abilities that are usually applied to humans such as intention and obligation (Weiss 1999).

Table 2.4 Types of Algorithms on Digital Platforms

	Not able to publish new content	**Able to publish new content**
Internal to the platform	Type I: Internal general-purpose algorithms	Type II: Internal machine agents
External to the platform	Type III: External general-purpose algorithms	Type IV: External machine agents

In addition, some algorithms are implemented by the platform as an integral part of platform infrastructure (e.g., product ranking based on clickstream data), whereas others are introduced by third parties to act on the focal platform (e.g., external bots and web scrapers). We define them as internal and external algorithms respectively. Many of the tasks these algorithms can perform may not be much different across four types. For example, all four types of algorithms can perform sorting, ranking, classification, and clustering tasks. Those that are able to generate new content will post the task outcomes on the platform in certain ways (e.g., publish new advertising information based on the ranking task). However, for the external algorithms, the platform may have some control mechanisms to prevent them from acquiring information and posting robot-generated content to avoid unintended consequences such as a data breach (e.g., Li et al. 2012, Thelwall and Stuart 2006).

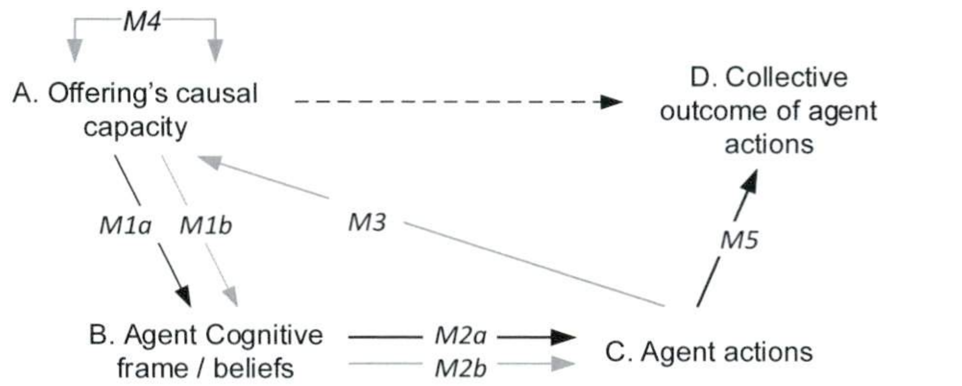
Human agents and machine agents share similar activities. Human agents make sense of the representations of offerings on the platform, generate various cognitions (or perceptions, beliefs, attitudes, desires), make decisions (e.g., evaluation, selection, comparison), and can post content and/or commit to the offering. Similarly, machine agents acquire information, build up knowledge about the environment in which they are situated, evaluate possible action plans, and can post content and/or commit to the offerings. The offerings do not act. They are posted by the providers, and their display is manipulated by the algorithms.

Two types of activities are of particular importance: commitment and non-committal actions. We define *commitment* actions as binding obligations that can take the form of a promise, payment, or bid related to the offering (e.g., crowdfunding contribution by human agents, auto-transaction by the machine agents). These are essentially money-related actions that accomplish the transaction. In contrast, *non-committal actions* are non-binding activities that can take the form of content creation, relationship building, or offering evaluation. Both human agents and machine agents can perform these actions.

The entities (i.e., human agents, algorithms and offerings) have temporospatial organizations that are usually constrained by platform rules and infrastructures. For example, the platform may constrain the speed of transaction, the sequence of interaction, and the accessibility of required information. In principle, human agents can interact with other human agents (e.g., make comments and replies) and machine agents (e.g., dialogue with a chatterbot), and machine agents can interact with other machine agents, either internal or external to the platform (e.g., web crawling activities may influence page ranking).

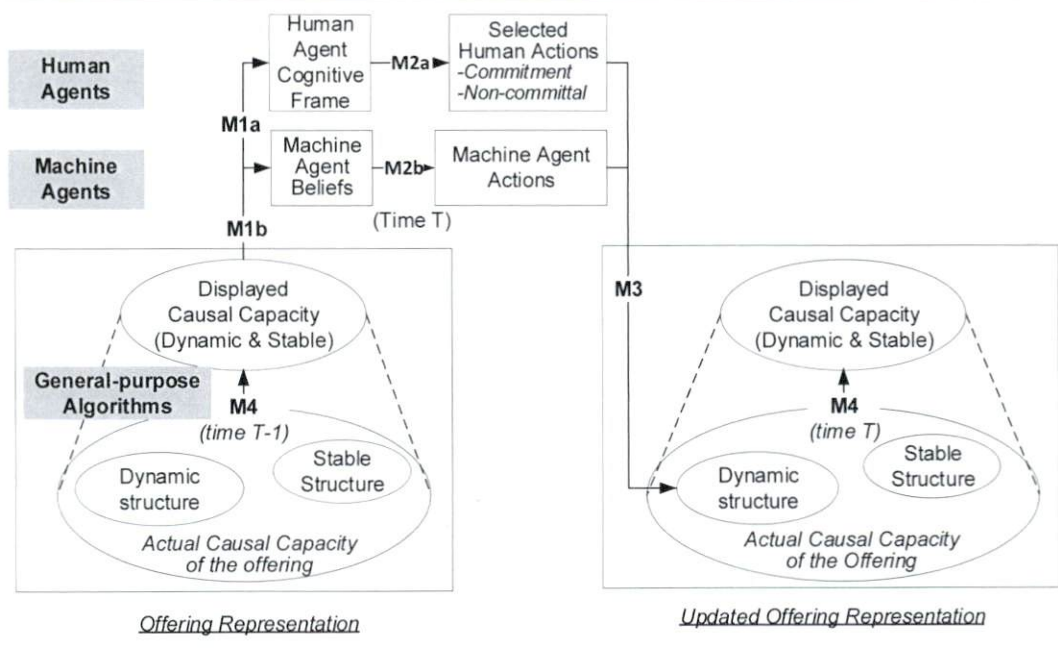
2.4.3 Constructing the Meta-Schema

Figure 2.2 is a high-level illustration of our meta-schema. It describes computational mechanisms (grey arrows) and social mechanisms (black arrows) that underlie the phenomenon (i.e., the relationships between platform offering representations, user commitment, and collective outcomes), which includes both constitutive explanation (i.e., causally relevant entities and activities) and causal processes (i.e., the enactment of agent actions over time). Figure 2.3 is a detailed representation of the meta-schema, and Table 2.5 describes each mechanism.



Note. The solid arrows represent mechanisms, not information flows, variance or causal effects between the constructs. The dotted arrow is a direct association, which is not a mechanism. The grey arrows represent computational mechanisms for machine agents, and the black arrows represent social mechanisms for human agents.

Figure 2.2 A Mechanism-Based Meta-Schema of the Impacts of Offering Representation on Digital Platforms



Note. The arrows represent mechanisms, not information flows, variance or causal effects between the constructs.

Figure 2.3 Detailed Representation of the Meta-Schema

Table 2.5 Summary of the Meta-Schema	
For Human Agents	**For Machine Agents**
M1a: human agent cognitive frame formation Human agents form cognitive frames (e.g., beliefs, attitudes, desires, and preferences) about the offering's causal capacity.	**M1b: machine agent belief formation** Machine agents collect percepts from the environment and build up knowledge of the environment.
M2a: human agent action formation Human agents' cognitive frames about the offering result in actions.	**M2b: machine agent action selection** Machine agents determine how to respond to events to achieve goals based on the knowledge about the environment (i.e., belief) and a collection of pre-defined plans, or they assemble plans from actions.
M3: causal capacity update Human agent and machine agent actions modify the values of the attributes of the dynamic structure of an offering.	
M4: offering representation update Internal general-purpose algorithms assess the values of the attributes of the offering causal structure and may modify the display of these values (e.g., ranking, labeling, and grouping) according to the goals of the platform.	
M5: collective outcome emergence Collective outcomes are produced by bottom-up coalescence (e.g., composition and compilation) of agent-level outcomes.	

2.4.3.1 Mechanism (1a) Human-Agent Cognitive Frame Formation

Human-agent cognitive frame formation is a situational mechanism that explains how social context influences agents' cognitive frames (e.g., beliefs, attitudes, perceptions, motivations, desires, preferences, affects[10]) that in turn constrain and enable agents' opportunities for action (Hedström and Swedberg 1998). As shown in Figures 2.2 and 2.3, an offering's displayed causal capacity, which is perceptible by human agents, influences their cognitive frames about the offering that in turn enable and constrain the opportunity for commitment and non-committal action.

[10] Despite the traditional affect-cognition distinction, neural mechanisms show that affect may be a type of cognition (Duncan and Barrett 2007). In this study, we take a broad view of cognition which includes affects.

This mechanism is in line with the Theory of Reasoned Action and the Theory of Planned Behavior (Ajzen 1985, Ajzen 1991, Fishbein and Ajzen 1975). Take human-agent belief formation as an example. If we stay at the behavioral and psychological level of explanation[11], human beliefs are "a person's subjective probability judgments concerning some discriminable aspect of the world" (Fishbein and Ajzen 1975 p.131). Beliefs reflect human agents' understanding of the environment –the offering, the digital platform environment, and the social environment on the digital platform. Direct observations, which provide descriptive information about objects (i.e., descriptive beliefs), and interactions, which provide information about unobservable characteristics (i.e., inferential beliefs), are two sources of belief formation (Fishbein and Ajzen 1975). For the offerings on digital platforms, the displayed causal structure, both the stable and the dynamic parts, provides individuals with opportunities to make sense of the object—to understand the causally relevant elements of an offering (e.g., the type, the description, past performance)—and to make links between the object and their beliefs about the attributes of the object. By interacting with the representations of these offering elements (e.g., virtual try-on, communication with providers), individuals can infer unobservable attributes of the offering elements (e.g., usability, trustworthiness). The formation of other cognitive and psychological concepts such as attitudes, desires and preferences shares much common ground with the formation of beliefs —there exists an object, and individuals need to make sense of the object first to establish the cognitive frames before advancing to decision-making and actions (Druckman and Lupia 2000, Moses et al. 2000, Regan and Fazio 1977). Thus, we propose that:

Proposition 1a: Human agents form cognitive frames (e.g., beliefs, attitudes, perceptions, motivations, desires, preferences, affects) through observational learning or direct interaction with the offering representation, which is a sense-making process to understand the offering's causal capacities.

The M1a mechanism is often implicit in variance-based explanations of user commitment on digital platforms. For example, the social media activity of a fundraiser is hypothesized

[11] Future research can go into lower-level mechanisms to explain belief formation, such as neuron activities in the brain.

to enhance social embeddedness of a crowdfunding campaign (i.e., the causal capacity of the offering), which in turn is thought to influence crowdfunding contributors' trust perceptions and social image concerns (Hong et al. 2018). Past sales information and eWOM on Groupon were found to influence subsequent incremental sales, the authors explaining that they serve as quality signals, updating customer beliefs via observational learning to increase their awareness and reduce the sense of quality uncertainty (Li and Wu 2018). Similarly, review volume and valence on eBay were found to influence buyers' willingness to pay through perceived risk and perceived value (Wu et al. 2013). From the evidence above, we can infer that human agents' cognitive frames have structures from which the formation of a belief or attitude may elicit subsequent beliefs and attitudes. For example, previous crowdfunding contributors that hide their identity information may elicit a sense of uncertainty about the crowdfunding campaign from future contributors, which reduces future contributors' confidence and increases skeptical perceptions (Burtch et al. 2016).

One approach to enhance precision and adequacy of explanation is to identify different complementary, substitutive, competing, or parallel elementary mechanisms at work that form the molecular mechanism (Hedström and Swedberg 1998). Extant research mostly relies on the signaling perspective which suggests that the presentation of the offering serves as a quality signal that induces human beliefs. A signaling mechanism can be molecular in that various elementary mechanisms may work simultaneously to produce cognitive dissonance, reinforce certain beliefs or create spillover effects. Hedström and Swedberg (1998) provided several prototypical elementary mechanisms which fit under this broad category of how the environment influences human agents' formation of cognitive frames (e.g., adaptive formation, wishful thinking, compensation, crowding out, contrasting, endowment). These elementary mechanisms can be applied to our context and are a good starting point upon which future research can construct a contextualized mechanism-based explanation.

2.4.3.2 Mechanism (1b) Machine-Agent Belief Formation

Similar to human belief formation, machine agents form *beliefs* before acting. Belief for a machine agent is defined as "the agent's knowledge or information about the environment, itself or other agents" (Padgham and Winikoff 2004 p.10). It is the cache of information (called *percepts*) that has been received from the environment, and the processing of these percepts is the sense-making process employed by machine agents. Take the PageRank algorithm from Google as an example. It measures the importance of webpages to order search results, thus allowing personalized display. Although the ranking results from PageRank are no longer available to the public since 2016, thereby making the actual ranking algorithm much more of a "secret", the algorithm itself represents a general framework for machine agents to perform ranking tasks through web graph analysis. PageRank collects information about the link structure (i.e., in-links and out-links) of a webpage, which serves as an indicator of the page value. The contribution of a link may be depreciated by its relevance, position (in the central place vs. at the bottom), topic or clickstreams, depending on how the algorithm defines the weights (Haveliwala 2003, Langville and Meyer 2011, Xing and Ghorbani 2004). If we consider a webpage as a node in the network graph, the collection of node attributes and network structure information (i.e., percepts) by the algorithm helps it makes sense of the environment and build up a weighted webpage network (i.e., belief formation), which decides the algorithm's subsequent ranking actions. Hence, we propose:

Proposition 1b: Machine agents' ability to build up knowledge of the environment (including knowledge of the offering and its surrounding social dynamics) varies according to their percepts collection and processing strategies.

For example, Russell and Norvig (2016) discussed four types of machine agents: (1) Simple reflex agents select actions based on current percepts, ignoring the past percept history, which limits the machine agents' knowledge of the environment to its current version. (2) Model-based reflex agents keep track of the percept history to reflect on the unobservable parts of the current percepts, allowing the agents' knowledge of the environment to continue updating. Under this strategy, machine agents are able to build up a historical view of the environment. (3) Goal-based agents require not only information about the environment, but also information about the goal, meaning that knowing the environment is not enough to decide whether or not the situation is desirable, and the goal is an integral part of machine agents' beliefs. Under this strategy, the resulting belief about the environment is contingent on the machine agents' acting to achieve the goal. (4) Utility-based agents, in addition to information about the environment and goal, follow a utility function to internalize performance measures about prospective behaviors in the environment. Therefore a machine agent builds up beliefs by comparing different environment states with the level of goal achievement.

Although traditionally considered a computer science topic, understanding machine-agent belief formation is relevant in the present context. On one hand, the percepts collected by machine agents are largely generated by human users. Consequently, the design of the machine agents should consider the situated cognition and actions of human agents. On the other hand, human agent belief formation is largely influenced by other algorithms present on the platform. For example, the format and content of recommendation agents (e.g., product ranking, personalized push messages) may be influenced by both human user profiles (e.g., users' past actions) and other machine agents' actions (e.g., be triggered by a virtual assistant, influenced by web crawlers' repetitive clicking actions).

2.4.3.3 Mechanism (2a) Human-Agent Action Formation

M2a is an action-formation mechanism, whereby agents' cognitive frames lead to or change the agents' behaviors (Hedström and Swedberg 1998). As per the cognitivist model assumption, individuals' behaviors toward an object largely depend on their cognitive frames about the object. This assumption underlies most extant variance-based studies which consider beliefs and attitudes as direct antecedents of actions. For example, trust and risk beliefs associated with the offering (e.g., toward the offering, toward the platform, toward the offering provider) are frequently examined antecedents of purchasing behaviors (e.g., Guo et al. 2018, Ou et al. 2014, Pavlou and Gefen 2004, Wu et al. 2013).

The composition of cognitive frames about an offering's causal capacity, however, is complex. Thus, the structure of individuals' cognitive frame can be difficult to detect, and identifying the right causal process can be challenging. Moreover, the offering's causal capacity is not limited to the offering itself but also includes the social dynamics that surround it. For example, if following Aristotle's modes of persuasion (Rapp 2010), the persuasive aspects of an offering, thus its causal capacity, can be derived from ethos (i.e., credibility of the presenter), pathos (i.e., the sentiment that appeals to the audience's emotion), and logos (i.e., the logical support). Therefore, individuals may not only evaluate the offering itself, but also the providers and how the offering is delivered on digital platforms. Such evaluation may be based on salient information that is explicitly presented on the platform, or on inferred information that individuals need to explore and connect in order to identify the probably true information. The presentation of the information can be static (e.g., the offering profile that is pre-designed) or dynamic (e.g., real-time peer activity information). Different types of information for evaluation may lead to different decision-making processes and actions (e.g., Jiang et al. 2018). Accordingly, human agents' cognitive frame compositions and the effort to deliver commitment versus non-committal actions can be different.

Even with the same cognitive frame composition, different cognitive frame formation patterns may result in different actions. For example, Hedström and Swedberg (1998) provided three classic examples based on similar beliefs but exhibiting different action formations. The first example is Robert Merton's (1968) self-fulfilling prophecy, which suggests that the initial false information about a situation will be reinforced during its diffusion, and will evoke behavior that eventually makes the false conceptions true. It is a cumulative and negative reinforcing process in which the wrong belief is enhanced though rumors and other people's wrong reactions. The second example is James Coleman's (1957) network diffusion process where an influential individual's opinion is diffused throughout the social network and influences what his or her friends say and do. The last example is Mark Granovetter's (1978) threshold theory, which posits that an individual's action differs as per the number of other actors who have done the same; thus there is a threshold in an individual's mind that guides his or her decision-making. All three examples involve beliefs regarding social validation of peer opinions, but different diffusion processes which lead to different belief formation patterns. These different patterns exhibit different impacts on an individual's propensity to act, and the impacts of beliefs on actions are far more complex than a linear process. Accordingly, we propose:

Proposition 2a: Human agents' cognitive frame compositions influence action selection, and different patterns of the cognitive formation may have different impacts on actions.

As has been discussed earlier, an individual's cognitive frame may include beliefs, attitudes, perceptions, desires, preferences, and any cognitive and psychological factors that can create opportunities for action. The evaluation of the offering causal capacity which creates opportunities for actions can be axiological in that the choice of actions may also depend on human agents' goals and expectations in the specific context.

2.4.3.4 Mechanism (2b) Machine-Agent Action Formation

Like human agents, machine agents act to affect the environment. After receiving or collecting information from the environment, a machine agent determines how to respond to achieve its goal, based on its belief about the environment and a collection of plans. The actions under this mechanism are mostly related to new content generation and direct interaction with human agents (e.g., analyzing user-generated data and responding accordingly). To illustrate, a recommendation agent may collect individuals' browsing histories to build consumer profiles and conduct matchmaking (e.g., content-based, social network-based, hybrid), so that it can provide personalized recommendations that influence individuals' purchasing decisions (Li and Karahanna 2015). A chatterbot may extract keywords from speech or input text to perform pattern matching (e.g., natural language enquiries, simple statements, semantic meaning enquiries) so that it can respond and converse with human users (Abdul-Kader and Woods 2015). The action performed by machine agents can be instantaneous or durational.

To form the action, a proactive agent pursues goals (i.e., if a plan fails to achieve the goal, the agent keeps trying the alternatives until it is no longer relevant), and a reactive agent responds to events that are significant occurrences, causing changes in information about the environment (i.e., percepts). As has been discussed earlier, an intelligent agent should be both proactive and reactive. To realize the goal, a machine agent needs a plan that specifies preconditions and effects, and can be pre-defined (i.e., plan library) or assembled from actions. The relationship between a machine agent's percepts, belief, goals, plans and actions can be very dynamic, yet rational. If an event occurs, it will be processed to update the machine agent's beliefs and may modify the (sub)goals (e.g., trigger new goals, drop impossible goals, change goal priorities). An action plan will be selected from the plan library. The execution of the action plan (i.e., enactment of the action) may yield new events, change (sub)goals, and modify beliefs (Padgham and Winikoff 2004).

Depending on the decision-making architecture, a machine agent has different strategies to decide which action to perform (i.e., action plan selection). For example, Weiss (1999) discussed four types of agent architectures: (1) a logic-based architecture in which the deduction rules (e.g., if-then function) and the database are encoded to derive the best action formulae using logical representation language; (2) a reactive architecture in which the best action is a product of agent-environment interaction (not a product of syntactic manipulation of the logical representations), thus the action should be situated in the environment and emerge from the interaction of various lower-level actions; (3) a belief-desire-intention architecture in which the best action is derived from deliberation and means-ends reasoning to achieve balance between multiple intentions; and (4) a layered architecture in which various subsystems are decomposed and arranged with hierarchies to deal with different types of goals and behaviors (e.g., one subsystem for proactive behaviors and another subsystem for reactive behaviors). Accordingly, we propose that:

Proposition 2b: Machine agents' action formation is influenced by their decision-making architecture (e.g., logic-deduction, reactive, belief-desire-intention, layered), so that even with the same percept and goals, the action plan selection may be different.

The purpose of discussing agent architectures and action formation mechanisms here is not to encourage everyone to conduct algorithm design research – indeed, many algorithm details are kept "secret" by a business since it is the intellectual assets that make the business competitive. However, the performative and material aspects of algorithms and machine agents are essential to understanding how and why *human agents* act, as well as understanding the changes on digital platforms (Orlikowski and Scott 2015). To illustrate, Gleasure et al. (2017) identified eight material features of book publishing crowdfunding platforms (e.g., reviews, reward, product description) and five material features of internally managed production activities (e.g., editing, curated mechanisms for uploading, physical premises), which are supported by algorithms and machine agents. The enactment of these algorithms creates social drivers (e.g., desires, pursue social identity, willingness) for human users to contribute to a campaign with different practices (e.g., a donation towards book vs. process) and motivations (e.g., the creation of impetus vs. creation of stress).

In summary, the enactment of machine agent actions contributes to the social dynamics on digital platforms, which shapes and constrains agent (both human and machine) belief and subsequent action formation.

2.4.3.5 Mechanism (3) Offering Causal Capacity Update

The new information created by human actions (i.e., commitment and non-committal) and machine agent actions (i.e., from both internal and external sources) may modify the values of the attributes of the offerings. For example, newly published reviews and ratings may enhance the attractiveness of the offering; consumer-provider communication information may enhance the trustworthiness of the provider; previous transaction listing may enhance the persuasiveness of the offering; and social media broadcasting presented on the platform may enhance the awareness of the brand. Such updated offering causal capacity will influence subsequent agent beliefs and actions. This updating process is a computational mechanism in which agent actions (both human and machine) generate data, and the platform's internal general-purpose algorithms receive, collect and process the new information to modify the offering's dynamic structure. The information processing execution can be selective depending on the objective of the platform and the decision-making architecture. For example, an algorithm following logical-deduction architecture retrieves the information processing plan from its encoded action formulae database, so the same type of information is always processed in the same way to update the offering's attributes. However, an algorithm following layered decision-making architecture may process the same piece of information differently, depending on the goal and environment (e.g., an AI algorithm learns from past user review examples and generate new policies for updating offering labels). As a result, different information sources (e.g., commitment action vs. non-committal actions, from internal machine agents vs. external machine agents) may yield different impacts on such update processes, depending on the algorithms' decision-making architecture. Hence, we propose:

Proposition 3: Offering causal capacity update is a joint process through which human and machine agent actions provide new information, and the platform's general-purpose algorithms selectively use this information to update the offering's attributes.

The computation process itself is most likely operating in the background, consequently non-observable by human agents, yet its enactment impacts agents. For example, cumulative sales indicate social validation of product quality, which means that each new transaction that occurs and is recorded on the digital platform updates the strength of an offering's quality signal. It should be noted that not all information generated from agent actions are salient and directly related to the offering value and quality. Various pieces of non-salient information can be generated during agent action (e.g., consumer browsing data, the expansion of product network), requiring extra effort in sense-making and connection building to ensure that causally relevant information is processed in appropriate ways to bring impact.

M3 is an important mechanism that explains how IT uses human artifacts (i.e., algorithms use human-generated information) and impacts human actions (Demetis and Lee 2018), which is different from traditional social mechanisms (e.g., M1a, M2a) that explain how humans use IT to act on digital platforms.

2.4.3.6 Mechanism (4) Offering Representation Update

M4 is the processing and display of the offering representation information on the digital platform in a way that is perceptible by human and machine agents. Not all newly available information is displayed on a platform; therefore, we differentiate the actual causal capacity and the displayed capacity that is perceptible by agents. It should be noted that offering representation is the description of offering attributes that are waiting to be interpreted, and displayed causal capacity refers to those attributes that are presented on the webpages. M4 is the process in which the offering representation is updated due to the change in the displayed attributes (i.e., displayed causal capacity) after selectively processing the actual causal capacity. For human agents, they mostly form their understanding based on displayed attributes on the screen; yet for machine agents, they have the ability to detect the information behind the screen (e.g., extract tags from HTML) which is also considered as a type of display in our study. In both cases, it is the update of offering representation that directly impacts subsequent agent belief and action formation.

After an offering's actual causal capacity is updated (M3), the platform may assess the values of the attributes of the offering and decide if and how to update the display of these attributes according to the goals of the platform. Again, the display selection process depends on the algorithms' decision-making architecture. For example, the platform may choose to exhibit a full transaction history instead of a cumulative sales index. It may choose to screen the negative reviews and only present the neutral or positive ones. Other potential operations may involve ranking, grouping, or labeling that not only screen and aggregate existing information but also give new meaning to these back-end data (e.g., an intelligent recommendation agent making new connections in a product network and generating novel recommendations based on learned customer behavior). However, as opposed to M2 and M3, no new data is generated by the algorithm in M4; it is the intelligent use of existing information that produces the impacts. Herein, we propose that:

Proposition 4: The platform's general-purpose algorithm selects information from the actual offering causal capacity to update the offering representation depending on its decision-making architecture.

2.4.3.7 Mechanism (5) Collective Outcome Emergence

Mechanism 5 is a transformational mechanism in which individual actors' actions produce social patterns and outcomes (e.g., the emergent macro-level properties) (Hedström and Swedberg 1998). It explains the emergence of collective outcomes from lower levels. Indeed, the outcomes of agents' commitments and non-committal actions coalesce at the collective level through value exchange, social interaction, and enactment processes. Such an emergence of collective outcomes is a bottom-up mechanism in which agent-level actions are realizers of the collective level outcomes.

Following multilevel research, we posit that the emergence process can take various forms, ranging from an isomorphic composition that linearly sums up similar parts to a discontinuous compilation that combines largely dissimilar parts in a nonlinear and patterned manner (Kozlowski and Klein 2000). Existing variance-based studies that modeled collective outcomes mainly take a composition approach, which assumes the collective outcomes are linear and incremental additions of individual contributions (e.g., Hong et al. 2018, Hu et al. 2017, Li and Wu 2018, Lin et al. 2017). Under this assumption, individual actions share common properties (e.g., same antecedents and processes) so that an average or sum can be used to understand the results of individual actions at the collective level. This may be because the types of collective outcomes that have been examined are mostly limited to the accumulation of monetary contributions (e.g., cumulative sales, sales ranks, funding progress), which is a linear addition process in which individuals' contributions are viewed as identical.

However, various collective outcomes cannot be linearly aggregated from individual contributions. For example, auctions usually result in taking the highest bid. The final interest rate for P2P lending can be a result of networked negotiation (i.e., multiple borrowers with multiple lenders for several similar lending requests). The reservation performance of a hotel can be compiled from several competing and complementary booking channels (e.g., directly visiting the platform, visiting the platform through social media recommendation, visiting the platform through Google recommendation). The overall subscription to a professional service (i.e., online medical consultation) can be a spillover result from related service adoption (e.g., use of offline medical service from a physician from the same hospital). Since compilation processes are less understood in the current literature, yet practically important, there remain great opportunities to understand collective outcomes at the offering or platform level by identifying patterns of individual actions and the interdependence between those patterns.

Proposition 5: The nature of individual commitment and non-committal actions will influence the emergence of collective outcomes, which will vary from composition (e.g., linear addition of individual transactions) to compilation (e.g., maximum bidding results, networked negotiation).

2.5 Discussion

2.5.1 Evaluating the Meta-Schema

We now turn to the assessment of our meta-schema. A mechanism schema may fail due to inappropriate characterization of the phenomenon, superficial modeling (i.e., only re-describing the phenomenon by showing correlations or causal relations that may be useful for prediction without describing the underlying mechanisms), incomplete construction (i.e., the mechanism sketches that have a black box for components for which their functional role is unknown), incorrectness (i.e., the schema fails to accurately describe the mechanism for the phenomenon) or empirical reasons such as experimental error, data analysis error, special cases, model anomalies, and falsifying anomalies (Craver and Darden 2013). Whereas the last issue can only be solved by empirical investigation, the three former issues can be improved by phenomenon recharacterization and iterative reconstruction.

To assess completeness, we examined the studies that were part of our literature review (see Appendix B). Although none applies a mechanism-based explanation, we identified mechanisms that were implicit in the authors' justification of the relationships they hypothesized. The majority of the studies imply M1a and M2a as their explanations. The elicited human agent belief towards the offering varies (e.g., social image concerns, awareness, quality uncertainty, perceived effort, perceived risk, perceived value), which contributes to commitment actions. Only two studies which examine collective outcomes discuss how these outcomes emerge (Gregg and Walczak 2008; Reiner 2014), but no studies directly hypothesize or examine this M5 mechanism. Furthermore, studies have not examined the role of machine agents and platform algorithms, thus have not considered the role of computational mechanisms. Although Gleasure et al.'s (2017) sociomaterial case study on the enacted material aspects of crowdfunding technology differentiates technological artefacts and social practices, it has not developed an explanation of how social practices are causally produced as a result of the manifestation of human-technology interaction.

In terms of correctness, our meta-schema involves a time effect such that previous actions—commitment or non-committal actions by human and machine agents—contribute to updating the offering causal capacity, which in turn influences subsequent agent actions. The accumulation of the result of agent-level actions gives rise to collective outcomes; therefore time is needed for the collective outcome to emerge. The sequence of activities (i.e., human agent actions, machine agent actions, and platform algorithms modifying the display of offering representations) is justified through forward and backward chaining, which confirms the correctness of the causal process.

A good schema should not only be complete and correct, but also possess some essential virtues if it is to inspire future research. To the best of our knowledge, there is no existing formal evaluation framework for assessing the validity and value of a mechanism-based explanation. Because we draw on Craver and Darden's (2013) mechanism discovery strategies, we also use their suggested criteria to evaluate the virtue of our meta-schema. The criteria can be classified into four types: (1) formal virtues that include testability and internal coherence, (2) pragmatic virtues that show how useful the theory is, such as fertility or conservation, (3) aesthetic virtues such as parsimony and elegance, and (4) empirical virtues such as accommodating well-established phenomena, having predictive power, being consistent with other non-rival theories, having generalizability, and unifying the diverse phenomena with a common pattern (Craver and Darden 2013). Table 2.6 synthesizes our self-assessment based on the first three criteria. As our work is theory development, the empirical virtues would be better assessed in future empirical studies.

Table 2.6 Meta-Schema Evaluation

Criteria	Description	Evaluating Our Meta-Schema
Testability	The extent to which the schema is testable with true observations (i.e., the mechanisms should reflect reality).	Future empirical studies should identify context-specific entities, activities and organizations, thus should be testable at a fine-grained level.
Internal coherence	The extent to which the schema is constructed in a coherent way with no contradictory elements.	The meta-schema does not have any contradictory elementary mechanisms.

Table 2.6 Meta-Schema Evaluation		
Fertility	The extent to which the schema can be used to develop new avenues of research (e.g., generate new research questions).	The meta-schema is a canvas on which future research can situate context-specific investigation. The new components of machine agents should generate new research questions that are practically and theoretically important.
Conservatism	The extent to which the schema retains crucial elements of the research tradition.	The meta-schema is in line with the research tradition in IS and marketing regarding the delivery of offerings, social impacts, and consumer commitment.
Parsimony	The extent to which the schema posits only those elements (e.g., entities, properties and relations) that are necessary.	With the key entities (human agents and algorithms including machine agents) and seven mechanisms, the meta-schema is a parsimonious representation of the phenomenon.
Elegance	The extent to which the schema effectively presents the elements so that they are organized in a compact manner.	By adopting the classic boat model, the meta-schema is compact and effective in explaining the phenomenon.

2.5.2 Leveraging the Meta-Schema to Develop Finer-Grained How-Plausible Mechanisms

By adopting our meta-schema as a canvas, future research can rely on the set of common vocabulary and prototypical mechanisms we propose to develop finer-grained how-plausible mechanisms with the support of context-specific empirical evidence. Here we discuss three approaches that we deem potentially useful: difference-making, agent-based modeling, and data-driven discovery. To illustrate the potential application of each approach, we use the vignette of a hypothetical online knowledge market where physicians provide medical consultation services via a multi-sided digital platform (e.g., Guo et al. 2017; Liu et al. 2016; Zhang et al. 2017). Figure 2.4 presents the essential elements of this platform. Here the phenomenon of interest is the relationship between the offering of free services for a limited amount of times by physicians (i.e., the freemium strategy) and the paid-service subscription conversion rate on the platform.

Platform Business Model Description

This hypothetical platform is based on real online medical consultation platforms such as HDF (HaoDaiFu), Ask A Doctor, and GoEVisit. The general business models across these platforms share many common characteristics. For example, the platform is an intermediary, connecting hundreds of thousands of physicians (or certified practitioners) with patients for non-emergency healthcare consultation. The platform provides various communication channels such as virtual conferencing, text-based service, and phone-based service. The matching between physicians and patients is usually realized by algorithms, ranging from simple assignment to dynamic profile matching. The platform usually has patient evaluation and rating mechanisms to indicate physician expertise and service quality. Many platforms adopt a freemium strategy—free trials are available for a limited number of times or periods to attract patients.

Phenomenon of Interest

The impact of freemium strategy on the conversion of paid services (i.e., subscribe to the paid- service after free-trials) on the platform

Key Components of the Mechanism (examples for illustration purposes)

Human Agent: physicians, patients

Machine Agent: physician-patient matching and recommendation systems

Commitment action (human): subscribe—or not—to the physician's medical consultation service

Non-committal action (human): review the physician, rate the physician after the service, free trials

Machine agent action: rank the physician profile page, provide a recommendation, automatic dialogue during the consultation (e.g., payment notification, patient health record posting)

Collective outcome: subscription conversion

Figure 2.4 A Hypothetical Digital Platform

Difference-making. The first potentially useful approach to reveal how the macro-level impact is brought about is difference-making, also called interventionist treatment, which follows counterfactual reasoning (Woodward 2011). It explains macro-level impacts by varying alternative micro-level actions and noting differences. The causal relevance is therefore established by showing covariational or contingency variables that play difference-making roles. For example, if the change in the magnitude of macro-level X is generally associated with the change in macro-level Y, and we wish to know more details about the possible mechanism M (at micro-level) that delivers such an effect, we may manipulate M (presence or absence, or different magnitudes) with controlled or natural experiments to examine whether or not the pattern of covariance can be found. Using the vignette example, a possible mechanism is that the free services offered by a socially validated "good" physician is highly valued compared with those offered by physicians with lower reputations. Thus, the supply and demand of highly-valued free services will give rise to platform-level service subscription conversion rates. Accordingly, natural experiments can be conducted in order to manipulate social validation processes and physician reputation delivery approaches to find the differential impacts. In addition to the traditional eWOM used to diffuse reputation, machine agents may play important roles—a recommendation agent may promote a physician based on patient eWOM and the algorithms' decision rules (e.g., physician response rate, service fee), or a matching agent may link a patient with a physician based on previous text-based consultation records. The difference-making manipulation can be done in various ways, and on both human and machines, to manipulate how social validation is delivered. If the manipulated components (e.g., reputation delivery approach) are found to be difference makers, the macro-level association becomes contingent on the physician's reputation, which is manifested in particular ways (e.g., based on algorithms rule or emergent from the patient eWOM diffusion network).

Agent-based simulation. Another potentially useful approach is agent-based simulation, which is considered to be a powerful tool for modeling the properties, activities and interactions among the components of a system (Marchionni and Ylikoski 2013, Smith and Conrey 2007). The simulation model is able to track the counterfactual dependencies so that we can make what-if inferences about how the organization of the components influences the behavior of the whole system (Fioretti 2012). For digital platform research, by simulating how human agents and machine agents act and interact, and by showing how the assumptions related to the offerings and actions make a difference, we can explain the bottom-up process by simulating the micro-level behavior rules and the relevant mechanisms. On the medical consultation platform, assumptions can be made regarding when a patient will pay—for instance, payment decision can depend on the type of disease, physician response rate, physician seniority, subscription fee, physician assignment by the algorithm, and so on – and the patient's characteristics (e.g., personality, communication style, health status) and platform rules (e.g., fixed amount of free offerings vs. flexible free offering by physicians) should also be taken into consideration. The parameter values and interaction rules set the foundation of a mechanism manifestation, and the plausibility of a mechanism is evaluated by running the computational program several times to assess the differential impacts of parameter inputs and the outcomes of interest.

Data-driven discovery. The third approach is inspired by the recent discussion of bridging theory-driven and data-driven research, partially due to the increasing availability of data and technologies that analyze and apply the massive amount of data (Abbasi et al. 2016; Agarwal and Dhar 2014; Rai 2016). Theory-driven research suggests that data should enter the theory construction process at the testing stage, so that researchers can start from a research problem, construct the hypothesis, and then be confronted with data. Data-driven research, on the other hand, begins with the actual data, so it is the data—rather than existing theory—that provides guidance for inquiry and new insight generation (Simon 1977). Maass et al. (2018) suggest that data-driven research and theory-driven research can be connected through two pathways: (1) patterns can be extracted from massive amounts of data (i.e., big data analytics) to develop or refine domain theory through abstraction and generalization; and (2) the domain theory can be used to identify data sources and types of analyses for further theory testing.

Both pathways are useful in conducting mechanism-based research. If the context-specific mechanism schema is available or can be potentially synthesized from literature, the schema can be used to determine the type of data needed (e.g., human agent action data) as well as the analysis approach for schema discovery. Then data-driven research can be employed to refine the initial mechanism schema. If the mechanism schema needs to be developed from scratch, the massive amount of data available on the platform can be used to detect an initial how-plausible schema, and more data requirements may emerge during the schema revision process as the entities and activities become more concrete. After multiple rounds of data pattern analysis, mechanism schema abstraction and schema revision, a finer-grained schema (i.e., the formal theory) will emerge. This is similar to a quantitative grounded theory approach (Glaser 2008; Walsh 2014) where explanations emerge from data patterns.

In our online medical consultation example, the platform hosts a massive amount of data regarding agent characteristics, agent interactions and commitment outcomes. Even if the "insider" data is difficult to obtain, displayed offering representation can be fetched by web scrapping. Text-based information can be processed using natural language processing, image data can be classified using machine learning image processing, and patterns in the numeric information can be detected using traditional data mining approaches or deep learning techniques if necessary. The objective is not to pursue the precision or accuracy of data analysis, but to discover possible mechanisms that can be further refined with a data-driven approach or other techniques such as difference-making and agent-based simulation.

In summary, to find finer-grained how-plausibly mechanisms, we need both probabilistic evidence (e.g., the association between the macro-level concepts such as social outcomes of certain social institutions) and lower-level mechanism-based evidence (e.g., the interactions between individuals) (Russo and Williamson 2007). To obtain mechanism-based evidence, various methods are available (e.g., difference-making, simulation, data-driven research). Combining mechanism-based thinking and statistical-based causal modeling is expected to be useful in discriminating competing mechanisms and identifying causality between the statistically connected explanans and explanandum.

2.6 Concluding Remarks

We propose a mechanism-based meta-schema to explain the relationships between representations of platform offerings, user commitment, and collective outcomes. The meta-schema decomposes the effect of a collective-level structure on individual-level entities, in our case, the machine agents and human agents. We theorize that both social mechanisms—which explain the actions of human agents—and computational mechanisms—which accomplish the functions of algorithms and machine agents—contribute to the realization of individual-level outcomes and the emergence of collective outcomes. In turn, individual-level outcomes coalesce to collective outcomes. For human agents, complex cognitive frames are formed from an offering's causal capacity, which consists of stable and dynamic attributes (M1a). These cognitive frames bring about action in a linear or nonlinear way (M2a). For machine agents, the platform is an environment where the agents collect percepts and build up knowledge (M1b). The machine agents act on the knowledge and plan to achieve functional goals (M2b). The results of agent actions update the offering's causal capacity by modifying the value of attributes of an offering (M3), and the platform algorithms assess the value of such an update and modify the display of these values (M4), which in turn influence agent perceptions of the offering. Thus, it is the offering's causal capacity and its dynamic update that connect the interaction of machine agents, human agents and offering representation. Over time, some collective outcomes at the offering or platform level may bottom out from agent action through composition or compilation processes.

Our study makes several contributions. First, in line with Rai and colleagues (2019), we re-conceptualize digital platforms at a collective entity where data are entered, screened, manipulated, displayed, and used, not only by human agents but also by machine agents and platform algorithms. Rather than ignoring or downplaying these latter entities, we explicitly theorize the mechanisms through which machine agents and platform algorithms, together with human agents, impact collective outcomes. These mechanisms offer a departure from prior literature by simultaneously considering human actions and machine actions, as well as their joint effects. Our work highlights how human actions on digital platforms can be shaped by algorithms, as well as how the user-generated data can be used by algorithms to exert power over future human actions (Demetis and Lee 2018).

Second, our meta-schema contributes to our understanding of digital platforms by proposing a mechanism-based explanation as an alternative or complementary theorizing approach to the dominant variance-based perspective. Whereas variance-based reasoning reveals the important antecedents associated with specific outcomes, a mechanism-based approach explicates *how* the outcomes are brought into effect through the manifestation of agent-level beliefs and actions. Our seven mechanisms reveal the cogs and wheels of how offering representation brings about user commitment and collective outcomes on digital platforms.

Finally, our meta-schema can provide a canvas for a significant body of future research on digital platforms. It offers a common vocabulary which forms the basis of an integrated understanding of the actions of machine agents, human agents, and algorithms. Our meta-schema also offers seven important mechanisms from which future researchers can develop their own finer-grained mechanisms for contextualized theories. Finally, we offer three examples of how to leverage our meta-schema – through difference-making, agent modeling, and data-driven discovery – in the hopes of mobilizing and inspiring a broad spectrum of future digital platform researchers.

References

Abbasi, A., Sarker, S., and Chiang, R. H. 2016. "Big Data Research in Information Systems: Toward an Inclusive Research Agenda," *Journal of the Association for Information Systems* (17:2), pp. I-XXXII.

Abdul-Kader, S. A., and Woods, J. 2015. "Survey on Chatbot Design Techniques in Speech Conversation Systems," *International Journal of Advanced Computer Science and Applications* (6:7), pp. 72-80.

Adomavicius, G., Bockstedt, J. C., Curley, S. P., and Zhang, J. 2013. "Do Recommender Systems Manipulate Consumer Preferences? A Study of Anchoring Effects," *Information Systems Research* (24:4), pp. 956-975.

Agarwal, R., and Dhar, V. 2014. "Editorial—Big Data, Data Science, and Analytics: The Opportunity and Challenge for Is Research," *Information Systems Research* (25:3), pp. 443-448.

Ajzen, I. 1985. "From Intentions to Actions: A Theory of Planned Behavior," in *Action Control*. Springer, pp. 11-39.

Ajzen, I. 1991. "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes* (50:2), pp. 179-211.

Andersen, H. 2012. "The Case for Regularity in Mechanistic Causal Explanation," *Synthese* (189:3), pp. 415-432.

Arnold, T., and Scheutz, M. 2018. "The "Big Red Button" Is Too Late: An Alternative Model for the Ethical Evaluation of Ai Systems," *Ethics and Information Technology* (20:1), pp. 59-69.

Avgerou, C. 2013. "Social Mechanisms for Causal Explanation in Social Theory Based Is Research," *Journal of the Association for Information Systems* (14:8), pp. 399-419.

Bechtel, W. 2009. "Explanation: Mechanism, Modularity, and Situated Cognition," in *The Cambridge Handbook of Situated Cognition*. Cambridge University Press, pp. 155-170.

Bechtel, W., and Abrahamsen, A. 2005. "Explanation: A Mechanist Alternative," *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* (36:2), pp. 421-441.

Benlian, A., Titah, R., and Hess, T. 2012. "Differential Effects of Provider Recommendations and Consumer Reviews in E-Commerce Transactions: An Experimental Study," *Journal of Management Information Systems* (29:1), pp. 237-272.

Bilgihan, A., and Bujisic, M. 2015. "The Effect of Website Features in Online Relationship Marketing: A Case of Online Hotel Booking," *Electronic Commerce Research and Applications* (14:4), pp. 222-232.

Bloemer, J. M., and Kasper, H. D. 1995. "The Complex Relationship between Consumer Satisfaction and Brand Loyalty," *Journal of Economic Psychology* (16:2), pp. 311-329.

Bogen, J. 2005. "Regularities and Causality; Generalizations and Causal Explanations," *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* (36:2), pp. 397-420.

Bunge, M. 2004. "How Does It Work? The Search for Explanatory Mechanisms," *Philosophy of the Social Sciences* (34:2), pp. 182-210.

Burtch, G., Ghose, A., and Wattal, S. 2013. "An Empirical Examination of the Antecedents and Consequences of Contribution Patterns in Crowd-Funded Markets," *Information Systems Research* (24:3), pp. 499-519.

Burtch, G., Ghose, A., and Wattal, S. 2016. "Secret Admirers: An Empirical Examination of Information Hiding and Contribution Dynamics in Online Crowdfunding," *Information Systems Research* (27:3), pp. 478-496.

Burton-Jones, A., McLean, E. R., and Monod, E. 2015. "Theoretical Perspectives in Is Research: From Variance and Process to Conceptual Latitude and Conceptual Fit," *European Journal of Information Systems* (24:6), pp. 664-679.

Carmi, E., Oestreicher-Singer, G., Stettner, U., and Sundararajan, A. 2017. "Is Oprah Contagious? The Depth of Diffusion of Demand Shocks in a Product Network," *Management Information Systems Quarterly* (41:1), pp. 207-221.

Coleman, J., Katz, E., and Menzel, H. 1957. "The Diffusion of an Innovation among Physicians," *Sociometry* (20:4), pp. 253-270.

Coleman, J. S. 1986. "Social Theory, Social Research, and a Theory of Action," *American Journal of Sociology* (91:6), pp. 1309-1335.

Constantinides, P., Henfridsson, O., and Parker, G. G. 2018. "Introduction-Platforms and Infrastructures in the Digital Age," *Information Systems Research* (29:2), pp. 381-400.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. 2009. *Introduction to Algorithms*. MIT press.

Craver, C. 2007. "Constitutive Explanatory Relevance," *Journal of Philosophical Research* (32), pp. 3-20.

Craver, C., and Tabery, J. 2017. "Mechanisms in Science," in: *The Stanford Encyclopedia of Philosophy,* E.N. Zalta (ed.). Metaphysics Research Lab, Stanford University.

Craver, C. F. 2001. "Role Functions, Mechanisms, and Hierarchy," *Philosophy of science* (68:1), pp. 53-74.

Craver, C. F., and Bechtel, W. 2007. "Top-Down Causation without Top-Down Causes," *Biology & Philosophy* (22:4), pp. 547-563.

Craver, C. F., and Darden, L. 2013. *In Search of Mechanisms: Discoveries across the Life Sciences*. University of Chicago Press.

Darden, L. 1991. *Theory Change in Science: Strategies from Mendelian Genetics*. Oxford University Press.

Darden, L. 2002. "Strategies for Discovering Mechanisms: Schema Instantiation, Modular Subassembly, Forward/Backward Chaining," *Philosophy of Science* (69:S3), pp. S354-S365.

Darden, L. 2008. "Thinking Again About Biological Mechanisms," *Philosophy of Science* (75:5), pp. 958-969.

Demetis, D. S., and Lee, A. S. 2018. "When Humans Using the IT Artifact Becomes IT Using the Human Artifact," *Journal of the Association for Information Systems* (19:10), pp. 929-952.

DistilNetworks. 2018. "2018 Bad Bot Report: The Year Bad Bots Went Mainstream." Retrieved March 6th, 2019, from https://resources.distilnetworks.com/whitepapers/2018-bad-bot-report.

Dodig-Crnkovic, G. 2011. "Significance of Models of Computation, from Turing Model to Natural Computation," *Minds and Machines* (21:2), pp. 301-322.

Dowe, P. 2010. "Causal Process Theories," in *The Oxford Handbook of Causation*. New York: Oxford University Press.

Druckman, J. N., and Lupia, A. 2000. "Preference Formation," *Annual Review of Political Science* (3:1), pp. 1-24.

Duncan, S., and Barrett, L. F. 2007. "Affect Is a Form of Cognition: A Neurobiological Analysis," *Cognition and Emotion* (21:6), pp. 1184-1211.

Eastlick, M. A., Lotz, S. L., and Warrington, P. 2006. "Understanding Online B-to-C Relationships: An Integrated Model of Privacy Concerns, Trust, and Commitment," *Journal of Business Research* (59:8), pp. 877-886.

Ellis, G. F. 2011. "Top-Down Causation and Emergence: Some Comments on Mechanisms," *Interface Focus* (2:1), pp. 126-140.

Elster, J. 1998. "A Plea for Mechanisms," in *Social Mechanisms: An Analytical Approach to Social Theory,* P. Hedström and R. Swedberg (eds.). New York: Cambridge University Press, pp. 45-73.

Evans, D. 2003. "Some Empirical Aspects of Multi-Sided Platform Industries," *Review of Network Economics* (2:3), pp. 1-19.

Feller, J., Gleasure, R., and Treacy, S. 2017. "Information Sharing and User Behavior in Internet-Enabled Peer-to-Peer Lending Systems: An Empirical Study," *Journal of Information Technology* (32:2), pp. 127-146.

Fioretti, G. 2012. "Agent-Based Simulation Models in Organization Science," *Organizational Research Methods* (16:2), pp. 227-242.

Fishbein, M., and Ajzen, I. 1975. *Belief, Attitude, Intention, and Behavior : An Introduction to Theory and Research*. Reading, Mass.: Addison-Wesley Pub. Co.

Ge, R., Feng, J., Gu, B., and Zhang, P. 2017. "Predicting and Deterring Default with Social Media Information in Peer-to-Peer Lending," *Journal of Management Information Systems* (34:2), pp. 401-424.

Ghoshal, A., Menon, S., and Sarkar, S. 2015. "Recommendations Using Information from Multiple Association Rules: A Probabilistic Approach," *Information Systems Research* (26:3), pp. 532-551.

Glaser, B. G. 2008. *Doing Quantitative Grounded Theory.* Sociology Press.

Gleasure, R., O'Reilly, P., and Cahalane, M. 2017. "Inclusive Technologies, Selective Traditions: A Socio-Material Case Study of Crowdfunded Book Publishing," *Journal of Information Technology* (32:4), pp. 326-343.

Glennan, S. 2002. "Rethinking Mechanistic Explanation," *Philosophy of Science* (69:S3), pp. S342-S353.

Glennan, S. 2010a. "Ephemeral Mechanisms and Historical Explanation," *Erkenntnis* (72:2), pp. 251-266.

Glennan, S. 2010b. "Mechanisms, Causes, and the Layered Model of the World," *Philosophy and Phenomenological Research* (81:2), pp. 362-381.

Glennan, S. 2011. "Singular and General Causal Relations: A Mechanist Perspective," in *Causality in the Sciences*. Oxford: Oxford University Press, pp. 789-817.

Glennan, S. S. 1996. "Mechanisms and the Nature of Causation," *Erkenntnis* (44:1), pp. 49-71.

Granovetter, M. 1978. "Threshold Models of Collective Behavior," *American Journal of Sociology* (83:6), pp. 1420-1443.

Gregg, D. G., and Walczak, S. 2008. "Dressing Your Online Auction Business for Success: An Experiment Comparing Two Ebay Businesses," *MIS Quarterly* (32:3), pp. 653-670.

Gu, B., Park, J., and Konana, P. 2012. "Research Note—the Impact of External Word-of-Mouth Sources on Retailer Sales of High-Involvement Products," *Information Systems Research* (23:1), pp. 182-196.

Guo, Y., Bao, Y., Stuart, B. J., and Le-Nguyen, K. 2018. "To Sell or Not to Sell: Exploring Sellers' Trust and Risk of Chargeback Fraud in Cross-Border Electronic Commerce," *Information Systems Journal* (28:2), pp. 359-383.

Guo, S., Guo, X., Fang, Y., and Vogel, D. 2017. "How Doctors Gain Social and Economic Returns in Online Health-Care Communities: A Professional Capital Perspective," *Journal of Management Information Systems* (34:2), pp. 487-519.

Halina, M. 2017. "Mechanistic Explanation and Its Limits," *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. Routledge, London, pp. 213-224.

Haveliwala, T. H. 2003. "Topic-Sensitive Pagerank: A Context-Sensitive Ranking Algorithm for Web Search," *IEEE Transactions on Knowledge and Data Engineering* (15:4), pp. 784-796.

Hedström, P., and Bearman, P. 2009. *The Oxford Handbook of Analytical Sociology*. Oxford University Press.

Hedström, P., and Swedberg, R. 1998. *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge University Press.

Hedström, P., and Ylikoski, P. 2010. "Causal Mechanisms in the Social Sciences," *Annual Review of Sociology* (36:1), pp. 49-67.

Hempel, C. G., and Oppenheim, P. 1948. "Studies in the Logic of Explanation," *Philosophy of Science* (15:2), pp. 135-175.

Hinz, O., Hill, S., Kim, J.-Y., Darmstadt, T. U., Karlsruhe Institute of, T., and Microsoft, R. 2016. "TV's Dirty Little Secret: The Negative Effect of Popular TV on Online Auction Sales," *MIS Quarterly* (40:3), pp. 623-644.

Ho, S. Y., and Bodoff, D. 2014. "The Effects of Web Personalization on User Attitude and Behavior: An Integration of the Elaboration Likelihood Model and Consumer Search Theory," *MIS Quarterly* (38:2), pp. 497-520.

Hong, Y., Hu;, Y., and Burtch, G. 2018. "Embeddedness, Prosociality, and Social Influence: Evidence from Online Crowdfunding," *MIS Quarterly* (42:4), pp. 1211-1224.

Hong, Y., and Pavlou, P. A. 2017. "On Buyer Selection of Service Providers in Online Outsourcing Platforms for IT Services," *Information Systems Research* (28:3), pp. 547-562.

Hu, N., Pavlou, P. A., and Zhang, J. 2017. "On Self-Selection Biases in Online Product Reviews," *MIS Quarterly* (41:2), pp. 449-A417.

Huang, Q., Chen, X. Y., Ou, C., Davison, R. M., and Hua, Z. S. 2017. "Understanding Buyers' Loyalty to a C2C Platform: The Roles of Social Capital, Satisfaction and Perceived Effectiveness of E-Commerce Institutional Mechanisms," *Information Systems Journal* (27:1), pp. 91-119.

Illari, P. 2013. "Mechanistic Explanation: Integrating the Ontic and Epistemic," *Erkenntnis* (78:2), pp. 237-255.

Illari, P., and Glennan, S. 2017. "Introduction: Mechanisms and Mechanical Philosophies," in *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. Routledge, pp. 19-28.

Illari, P. M., and Williamson, J. 2011. "Mechanisms Are Real and Local," in *Causality in the Sciences,* P. Illari, *et al.* (eds.). Oxford: Oxford University Press, pp. 818-844.

Illari, P. M., and Williamson, J. 2012. "What Is a Mechanism? Thinking About Mechanisms across the Sciences," *European Journal for Philosophy of Science* (2:1), pp. 119-135.

Jaccard, J., and Jacoby, J. 2009. *Theory Construction and Model-Building Skills: A Practical Guide for Social Scientists*. Guilford Press.

Jiang, Y., Ho, Y.C., Yan, X., and Tan, Y. 2018. "Investor Platform Choice: Herding, Platform Attributes, and Regulations," *Journal of Management Information Systems* (35:1), pp. 86-116.

Jones, M. L. 2018. "Silencing Bad Bots: Global, Legal and Political Questions for Mean Machine Communication," *Communication Law and Policy* (23:2), pp. 159-195.

Kim, M.-S., and Ahn, J.-H. 2007. "Management of Trust in the E-Marketplace: The Role of the Buyer's Experience in Building Trust," *Journal of Information Technology* (22:2), pp. 119-132.

Kozlowski, S. W. J., and Klein, K. J. 2000. "A Multilevel Approach to Theory and Research in Organizations Contextual, Temporal, and Emergent Processes," in *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions,* K.J. Klein and S.W.J. Kozlowski (eds.). San Francisco: Jossey-Bass, pp. 3-90.

Kuan, K. K. Y., Zhong, Y., and Chau, P. Y. K. 2014. "Informational and Normative Social Influence in Group-Buying: Evidence from Self-Reported and EEG Data," *Journal of Management Information Systems* (30:4), pp. 151-178.

Kuorikoski, J. 2012. "Mechanisms, Modularity and Constitutive Explanation," *Erkenntnis* (77:3), pp. 361-380.

Langville, A. N., and Meyer, C. D. 2011. *Google's Pagerank and Beyond: The Science of Search Engine Rankings*. Princeton University Press.

Levina, N., and Arriaga, M. 2014. "Distinction and Status Production on User-Generated Content Platforms: Using Bourdieu's Theory of Cultural Production to Understand Social Dynamics in Online Fields," *Information Systems Research* (25:3), pp. 468-488.

Levy, A., and Bechtel, W. 2013. "Abstraction and the Organization of Mechanisms," *Philosophy of Science* (80:2), pp. 241-261.

Li, S. S., and Karahanna, E. 2015. "Online Recommendation Systems in a B2C E-Commerce Context: A Review and Future Directions," *Journal of the Association for Information Systems* (16:2), pp. 72-107.

Li, X., and Wu, L. 2018. "Herding and Social Media Word-of-Mouth: Evidence from Groupon," *MIS Quarterly* (42:4), pp. 1331-1351.

Li, Z., Zhang, K., Xie, Y., Yu, F., and Wang, X. 2012. "Knowing Your Enemy: Understanding and Detecting Malicious Web Advertising," *Proceedings of the*

2012 ACM conference on Computer and communications security: ACM, pp. 674-686.

Lin, Z., Goh, K.-Y., and Heng, C.-S. 2017. "The Demand Effects of Product Recommendation Networks: An Empirical Analysis of Network Diversity and Stability," *MIS Quarterly* (41:2), pp. 397-A310.

Little, D. 2011. "Causal Mechanisms in the Social Realm," in *Causality in the Sciences,* P. Illari, *et al.* (eds.). Oxford: Oxford University Press, pp. 273-295.

Liu, X., Guo, X., Wu, H., and Wu, T. 2016. "The Impact of Individual and Organizational Reputation on Physicians' Appointments Online," *International Journal of Electronic Commerce* (20:4), pp. 551-577.

Maass, W., Parsons, J., Purao, S., Storey, V. C., and Woo, C. 2018. "Data-Driven Meets Theory-Driven Research in the Era of Big Data: Opportunities and Challenges for Information Systems Research," *Journal of the Association for Information Systems* (19:12), pp. 1253-1273.

Machamer, P., Darden, L., and Craver, C. F. 2000. "Thinking About Mechanisms," *Philosophy of Science* (67:1), pp. 1-25.

Manzo, G. 2006. "Generative Mechanisms and Multivariate Statistical Analysis: Modeling Educational Opportunity Inequality with a Multi-Matrix Log-Linear Topological Model: Contributions and Limitations," *Quality and Quantity* (40:5), pp. 721-758.

Marchionni, C., and Ylikoski, P. 2013. "Generative Explanation and Individualism in Agent-Based Simulation," *Philosophy of the Social Sciences* (43:3), pp. 323-340.

Merton, R. K. 1968. *Social Theory and Social Structure*. New York: Free Press .

Meseguer, J. 1992. "Conditional Rewriting Logic as a Unified Model of Concurrency," *Theoretical Computer Science* (96:1), pp. 73-155.

Miłkowski, M. 2011. "Beyond Formal Structure: A Mechanistic Perspective on Computation and Implementation," *Journal of Cognitive Science* (12), pp. 359-379.

Miłkowski, M. 2014. "Computational Mechanisms and Models of Computation," *Philosophia Scientiæ. Travaux d'histoire et de philosophie des sciences* (18:3), pp. 215-228.

Mithas, S., and Krishnan, M. S. 2009. "From Association to Causation Via a Potential Outcomes Approach," *Information Systems Research* (20:2), pp. 295-313.

Mohr, L. 1982. "Approaches to Explanations & Variance Theory and Process Theory," in *Explaining Organizational Behavior*. San Francisco: Jossey-Bass, pp. 35-70.

Moses, L., Coon, J., and Wusinich, N. 2000. "Young Children's Understanding of Desire Formation," *Developmental psychology* (36:1), pp. 77-90.

NarrativeScience. 2018. "Outlook on Artificial Intelligence in the Enterprise." Retrieved March 6th, 2019, from https://narrativescience.com/research-report/research-report-outlook-on-ai-for-the-enterprise/

Neff, G., and Nagy, P. 2016. "Talking to Bots: Symbiotic Agency and the Case of Tay," *International Journal of Communication* (10), pp. 4915-4931.

Oestreicher-Singer, G., and Sundararajan, A. 2012. "Recommendation Networks and the Long Tail of Electronic Commerce," *MIS Quarterly* (36:1), pp. 65-83.

Oliver, R. L. 1999. "Whence Consumer Loyalty?," *Journal of marketing* (63:4), pp. 33-44.

Orlikowski, W. J., and Scott, S. V. 2015. "The Algorithm and the Crowd: Considering the Materiality of Service Innovation," *MIS Quarterly* (39:1), pp. 201-216.

Ou, C. X., Pavlou, P. A., Davison, R. M. 2014. "Swift Guanxi in Online Marketplaces: The Role of Computer-Mediated Communication Technologies," *MIS Quarterly* (38:1), pp. 209-230.

Padgham, L., and Winikoff, M. 2004. *Developing Intelligent Agent Systems : A Practical Guide*. Chichester, England ;: John Wiley.

Pavlou, P. A., and Gefen, D. 2004. "Building Effective Online Marketplaces with Institution-Based Trust," *Information Systems Research* (15:1), pp. 37-59.

Piccinini, G. 2007. "Computing Mechanisms," *Philosophy of Science* (74:4), pp. 501-526.

Piccinini, G. 2017. "Computation in Physical Systems," in: *The Stanford Encyclopedia of Philosophy,* E.N. Zalta (ed.). Metaphysics Research Lab, Stanford University.

Piccinini, G., and Scarantino, A. 2010. "Computation Vs. Information Processing: Why Their Difference Matters to Cognitive Science," *Studies in History and Philosophy of Science Part A* (41:3), pp. 237-246.

Povich, M., and Craver, C. F. 2018. "Mechanistic Levels, Reduction, and Emergence," *The Routledge Handbook of Mechanisms and Mechanical Philosophy*, pp. 185-197.

Psillos, S. 2011. "The Idea of Mechanism," in *Causality in Sciences,* P. Illari, *et al.* (eds.). Oxford: Oxford University Press, pp. 771-788.

Putnam, H. 1991. *Representation and Reality*. MIT press.

Rai, A. 2016. "Editor's Comments: Synergies between Big Data and Theory," *MIS quarterly* (40:2), pp. iii-ix.

Rai, A., Constantinides, P., and Sarker, S. 2019. "Editor's Comments: Next-Generation Digital Platforms: Toward Human–AI Hybrids," *MIS Quarterly* (43:1), pp. iii-ix.

Rapp, C. 2010. "Aristotle's Rhetoric," in: *The Stanford Encyclopedia of Philosophy,* E.N. Zalta (ed.). Metaphysics Research Lab, Stanford University.

Recode. 2018. "Facebook Has Disabled Almost 1.3 Billion Fake Accounts over the Past Six Months." Retrieved March 29th, 2019, from https://www.recode.net/2018/5/15/17349790/facebook-mark-zuckerberg-fake-accounts-content-policy-update

Regan, D. T., and Fazio, R. 1977. "On the Consistency between Attitudes and Behavior: Look to the Method of Attitude Formation," *Journal of Experimental Social Psychology* (13:1), pp. 28-45.

Reiner, J., Natter, M., and Skiera, B. 2014. "The Impact of Buy-Now Features in Pay-Per-Bid Auctions," *Journal of Management Information Systems* (31:2), pp. 77-104.

Rivard, S. 2014. "The Ions of Theory Construction," *MIS Quarterly* (38:2), pp. iii-xiv.

Russell, S. J., and Norvig, P. 2016. *Artificial Intelligence: A Modern Approach*. Malaysia; Pearson Education Limited.

Russo, F., and Williamson, J. 2007. "Interpreting Causality in the Health Sciences," *International Studies in the Philosophy of Science* (21:2), pp. 157-170.

Salmon, W. C. 1984a. *Scientific Explanation and the Causal Structure of the World*. Princeton University Press.

Salmon, W. C. 1984b. "Scientific Explanation: Three Basic Conceptions," *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*: Philosophy of Science Association, pp. 293-305.

Shagrir, O. 2006. "Why We View the Brain as a Computer," *Synthese* (153:3), pp. 393-416.

Siegelmann, H. T., and Sontag, E. D. 1994. "Analog Computation Via Neural Networks," *Theoretical Computer Science* (131:2), pp. 331-360.

Simon, H. A. 1977. *Models of Discovery : and Other Topics in the Methods of Science*. Dordrecht, Holland ;: D. Reidel Pub. Co.

Smith, E. R., and Conrey, F. R. 2007. "Agent-Based Modeling: A New Approach for Theory Building in Social Psychology," *Personality and Social Psychology Review* (11:1), pp. 87-104.

Stinchcombe, A. L. 1991. "The Conditions of Fruitfulness of Theorizing About Mechanisms in Social Science," *Philosophy of the Social Sciences* (21:3), pp. 367-388.

Thelwall, M., and Stuart, D. 2006. "Web Crawling Ethics Revisited: Cost, Privacy, and Denial of Service," *Journal of the American Society for Information Science and Technology* (57:13), pp. 1771-1779.

Thies, F., Wessel, M., and Benlian, A. 2016. "Effects of Social Interaction Dynamics on Platforms," *Journal of Management Information Systems* (33:3), pp. 843-873.

Thies, F., Wessel, M., and Benlian, A. 2018. "Network Effects on Crowdfunding Platforms: Exploring the Implications of Relaxing Input Control," *Information Systems Journal* (28:6), pp.1239-1262.

Tiwana, A., Konsynski, B., and Bush, A. A. 2010. "Research Commentary—Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics," *Information Systems Research* (21:4), pp. 675-687.

Van de Ven, A. H. 2005. "Alternative Approaches for Studying Organizational Change," *Organization Studies* (26:9), pp. 1377-1404.

Victor, V., Thoppan, J. J., Nathan, R. J., and Maria, F. F. 2018. "Factors Influencing Consumer Behavior and Prospective Purchase Decisions in a Dynamic Pricing Environment—an Exploratory Factor Analysis Approach," *Social Sciences* (7:9), pp. 1-14.

Viglia, G., Pera, R., and Bigne, E. 2018. "The Determinants of Stakeholder Engagement in Digital Platforms," *Journal of Business Research* (89), pp. 404-410.

Wang, J.-J., Wang, L.-Y., and Wang, M.-M. 2018. "Understanding the Effects of eWOM Social Ties on Purchase Intentions: A Moderated Mediation Investigation," *Electronic Commerce Research and Applications* (28), pp. 54-62.

Walsh, I. 2014. "Using Quantitative Data in Mixed-Design Grounded Theory Studies: An Enhanced Path to Formal Grounded Theory in Information Systems," *European Journal of Information Systems* (24:5), pp. 531-557.

Weiss, G. 1999. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge: MIT Press.

Wimsatt, W. C. 1997. "Aggregativity: Reductive Heuristics for Finding Emergence," *Philosophy of Science* (64), pp. S372-S384.

Woodward, J. 2002. "What Is a Mechanism? A Counterfactual Account," *Philosophy of Science* (69:S3), pp. S366-S377.

Woodward, J. 2011. "Mechanisms Revisited," *Synthese* (183:3), pp. 409-427.

Woodward, J. 2017. "Scientific Explanation," in: *The Stanford Encyclopedia of Philosophy,* E.N. Zalta (ed.). Metaphysics Research Lab, Stanford University.

Wooldridge, M. 2009. *An Introduction to Multiagent Systems*. John Wiley & Sons.

Wooldridge, M., and Jennings, N. R. 1995. "Intelligent Agents: Theory and Practice," *The Knowledge Engineering Review* (10:2), pp. 115-152.

Wright, C. D. 2012. "Mechanistic Explanation without the Ontic Conception," *European Journal for Philosophy of Science* (2:3), pp. 375-394.

Wu, J., Arturo, E., and Gaytán, A. 2013. "The Role of Online Seller Reviews and Product Price on Buyers' Willingness-to-Pay: A Risk Perspective," *European Journal of Information Systems* (22:4), pp. 416-433.

Xing, W., and Ghorbani, A. 2004. "Weighted Pagerank Algorithm," *Proceedings. Second Annual Conference on Communication Networks and Services Research, 2004.*: IEEE, pp. 305-314.

Xu, J. J., and Chau, M. 2018. "Cheap Talk? The Impact of Lender-Borrower Communication on Peer-to-Peer Lending Outcomes," *Journal of Management Information Systems* (35:1), pp. 53-85.

Ylikoski, P. 2013. "Causal and Constitutive Explanation Compared," *Erkenntnis* (78:S2), pp. 277-297.

Ylikoski, P. K. 2017. "Social Mechanisms," in *The Routledge Handbook of Mechanisms and Mechanical Philosophy,* S. Glennan and P. Illari (eds.). London and New York: Routledge, pp. 401-412.

Zednik, C. 2017. "Mechanisms in Cognitive Science," in *The Routledge Handbook of Mechanisms and Mechanical Philosophy,* S. Glennan and P. Illari (eds.). London and New York: Routledge, pp. 389-400.

Zhang, X., Guo, X., Lai, K.-h., Yin, C., and Meng, F. 2017. "From Offline Healthcare to Online Health Services: The Role of Offline Healthcare Satisfaction and Habits," *Journal of Electronic Commerce Research* (18:2), pp. 138-154.

Zheng, H., Xu, B., Zhang, M., and Wang, T. 2018. "Sponsor's Cocreation and Psychological Ownership in Reward-Based Crowdfunding," *Information Systems Journal* (28:6), pp. 1213-1238.

Chapter 3 - Essay 3

Which Physicians Attract Paying Customers? Mining Massive Platform Data to Understand Patient Payment in Freemium-Based Online Medical Consultation

Jinglu Jiang
Department of Information Technologies, HEC Montreal
3000, chemin Côte-Sainte-Catherine
Montréal, QC
H3T 2A7
CANADA
jinglu.jiang@hec.ca

Ann-Frances Cameron
Department of Information Technologies, HEC Montreal
3000, chemin Côte-Sainte-Catherine
Montréal, QC
H3T 2A7
CANADA
ann-frances.cameron@hec.ca

Abstract

Online medical consultation as an emerging type of digital healthcare service has attracted much attention in recent years. Due to the nature of online consultation for health issues, it can be difficult for patients to accept such services and pay for them. Moreover, healthcare resource distribution on the platform may be highly unequal, allowing patients to be aware of only a small amount of healthcare services. Freemium, a multi-tier pricing model combining free trials and premium services, can be a potentially useful business model for attracting a larger user base and increasing service acceptance. However, the best ways to promote premium payment and help the platform identify high-value services and service providers in such a context are largely unknown. In this study, we explore the key online medical consultation service features that are associated with premium payment (as opposed to free-trial-only services). We use a massive dataset from an online medical consultation platform that has already achieved some success with a freemium model. Eight machine learning algorithms are used to cross-validate the model (based on machine learning performance measures) and the importance of the selected features (based on feature importance scores). The results show that, as compared to factors related to physician reputation, service-related factors such as service delivery quality (e.g., consultation dialogue intensity, physician response rate), patient source (e.g., online versus offline returning patients) and patient involvement (e.g., provide social returns, reveal previous treatment) appear to contribute more to premium payment.

Keywords: online medical consultation, digital healthcare service, digital platform, machine learning, classification, healthcare service quality, physician reputation

3.1 Introduction

A new wave of healthcare digitization is underway, and digital health companies are facing enormous opportunities and challenges to transform the healthcare industry in various ways. In the US, the health technology sector is forecast to reach $280 billion by 2021, and virtual health is expected to be a key solution for healthcare accessibility and long-term patient-provider collaboration (Deloitte 2019). In China, as of July 2018, the market share for the ICT-based health services was about $4.25 billion, with expected growth to 5.96 billion by 2019[12]. A McKinsey healthcare consumer survey (2018) reports that more than 70% of patients prefer digital healthcare solutions such as doctor search, health metrics self-monitoring and digital appointment than in-person or phone-based services. The increasing involvement of the private sector in recent years brings new digital products, healthcare channels, and business models that are not only helping organizations to be more efficient in healthcare service delivery but also influencing individuals' decision making and their accessibility to a broader range of healthcare choices (Biesdorf and Niedermann 2014).

The current study focuses on an emerging type of digital healthcare service – online medical consultation. It generally takes three forms: (1) an online community or social media model that focuses on peer-based Q&As and healthcare information sharing (e.g., communities like PatientsLikeMe and MedHelp); (2) a telemedicine model where the digital platform (as a third party) collaborates with designated healthcare professionals or healthcare organizations to deliver paid consultation services online (e.g., OnCall); and (3) a multi-sided market to which qualified healthcare professionals can register in order to provide various services (e.g., Practo, DocApp, EClinic247). This study focuses on the third form of consultation, which is delivered on multi-sided digital platforms, connecting healthcare providers and patients (or their caregivers). Traditionally, offering such professional services to individuals was mostly constrained by expert resources being available at a specific geographic location – like going through a specific law firm to find

[12] This market includes health portals, health ecosystems, online communities, online medical consultations, and mobile-based services such as health information apps, mHealth devices, and drug-choice apps. Accessed from https://www.analysys.cn/article/analysis/detail/20018737.

a lawyer, or hiring a registered accountant through an accounting agency. The emergence of digital platform-based consultation is considered a promising solution to healthcare resource shortages (especially for the remote areas) and inefficient healthcare resource distribution (Liu et al. 2016). Various forms of online medical consultation exist, such as AI-based chatbots (e.g., Babylon) and videoconferencing-based applications that connect physician and patients, which are usually implemented by healthcare institutions. In the current study, we only focus on human-delivered healthcare consultation services through third-party multisided digital platforms.

Online medical consultation is now being offered by many healthcare professionals around the world, and the market is projected to proliferate in the next five years (GarnerInsights 2018). Various benefits have been suggested such as healthcare resource accessibility, time savings for both doctors and patients, and cost savings (Chada 2017; Greenhalgh et al. 2016; Jiang, 2019; Shaw et al. 2018). From a business perspective, the involvement of healthcare companies and digital platforms creates an increasingly dynamic healthcare delivery ecosystem. However, areas of concern exist which may impede patients' acceptance or the establishment of long-term and repetitive interaction with the physician online. For example, clinical risk, regulatory challenges and healthcare equality issues may raise overall concerns about this new healthcare delivery approach (CBC 2017; Greenhalgh et al. 2016). Dissatisfaction may occur after an initial failed experience, and the diagnoses or treatment recommendations may be made with limited consideration of patients' medical history (Singh et al. 2018). Not surprisingly, many of these platforms are struggling to attract and retain patients who are willing to pay. In addition, the online medical consultation market mostly follows the Pareto principle in that 80% of the services are provided by 20% of the physicians on the platform (Li et al., 2016). Consequently, healthcare service providers on the platform have the additional challenge of standing out in a crowd of physicians who can provide comparable services (Cao et al. 2017; Li et al. 2019a; Yang et al. 2015).

Previous studies on online medical consultation have investigated various individual and contextual factors associated with patient selection of physicians and payment that contribute to a physician's success on the platform. For example, physician reputation, pricing, online rating and service quality perception are factors that are frequently examined (Deng et al. 2019; Sun et al. 2019; Wu and Lu 2018). While these studies provide useful managerial implications, they examine a small set of factors in isolation. In addition, these studies do not take into account the specifics of the business model – an important context that defines the value propositions and revenue generation mechanisms of the digital platform. Understanding a patient's service selection and payment decision under a specific business model is important, because the business model decides the platform's offering distribution (e.g., who should provide what service to which patient segment), the delivery of technological functionalities, pricing structures and more (Chesbrough 2010). Additional research is needed to develop a more holistic understanding of the factors that contribute to patient selection of physicians and payment, as well as an understanding that takes into account the characteristics of the digital platform's business model.

As a first step, this study focuses on online medical digital platforms employing freemium (FREE+ preMIUM), a business model with a multi-tiered pricing strategy that allows the coexistence of free (i.e., free trials) and paid (i.e., premium) versions of goods and services (Kumar 2014; Wagner et al. 2014). It seems to be a promising model that may help solve practical issues currently facing the online medical consultation market, such as patient reluctance and healthcare resource distribution inequality. Indeed, some medical consultation platforms have implemented a freemium model and achieved success. For example, PingAn Good Doctor won Hong Kong IPO approval in 2018 and posted a revenue of 1.12 billion from January to June 2018, and WeDoctor, a giant platform used by more than 2,700 hospitals in China with 160 million registered users, is valued at $5.5 billion pre-IPO[13].

[13] Reported by Bloomberg and Reuters: (1) https://www.bloomberg.com/news/articles/2018-07-02/a-6-billion-china-startup-wants-to-be-the-amazon-of-health-care; (2) https://www.reuters.com/article/us-

In other fields of online businesses, the freemium model has been adopted for digital products such as software and music streaming (Kumar 2014). For example, as of 2017, in-app purchase, rather than paid-download or upfront purchase, is used in 79% of the gaming apps and 50% of the non-gaming apps in the Google Play Store (Statista 2018). Other examples of successful online companies that employ a freemium model include the professional networking service LinkedIn, the digital file hosting service Dropbox and the online music streaming service Spotify. The success of various digital companies makes freemium a promising revenue-generation strategy, especially for digital content providers to make money from consumers who used to have a "for free" mentality (Wagner et al. 2014). Freemium usually acts as a marketing tool to attract a large user base and speed up product diffusion. It is especially useful for businesses providing experience or credence goods because free trials allow consumers to directly experience the products or services so that they can make a cost-free evaluation before purchase. However, converting free-trial consumers to premium subscribers can be difficult, and the free trials may cannibalize the adoption of the paid version (Bawa and Shoemaker 2004). Not surprisingly, there are a large number of freemium businesses which become trapped by the cost of offering free products and fail to attract enough premium payments.

Since online medical consultation is a type of online business that provides digital content, adopting a freemium strategy may be helpful to expand a user base that is more likely to develop a strong commitment to the platform, thus paying in future. The freemium model makes a lot of sense in this context because many of the practical issues are largely related to the nature of medical consultation (e.g., high patient uncertainty and lack of trust) and governance inefficiencies such as lack of action visibility and a poor matching process between physicians and patients (Hansen et al. 2019; Yellowlees et al. 2015). Thus, it is difficult for a patient to accept this kind of professional service online and establish long-term virtual collaboration. Because medical consultation is a type of credence good, and consumer uncertainty or psychological cost can be very high (Dulleck and Kerschbamer 2006; Nelson 1970), allowing patients to explore the service for free can be a useful

china-tencent-wedoctor/tencents-wedoctor-raises-500-million-values-firm-at-5-5-billion-pre-ipo-idUSKBN1IA08G approach to alleviate these concerns. Moreover, freemium can be a potential solution to better arrange and distribute professional resources. For example, the platform can focus more on promoting physicians who are not popular but have attributes similar to "top" physicians and can try to increase the awareness of their free service offerings among patients. Consequently, patients will have more service options and select services from a broader range of physicians due to the initial experience of free trials. This process may partially alleviate sparse service delivery related to the 80/20 problem as well as the issue of healthcare inequality.

Although empirical data from freemium-based platforms has been frequently used (e.g., Cao et al. 2017; Deng et al. 2019; Guo et al. 2017), to date few studies have explicitly acknowledged the freemium model when investigating online medical consultation (Liu et al. 2018 is one notable exception). In addition, previous freemium literature may not fully apply, given the unique characteristics of online medical consultation as a knowledge-intensive and highly specialized type of service. More research is needed to understand important service features that contribute to patient selection of physicians and payment by taking into account the nature of the freemium business model.

The objective of this study is to investigate this research need. Although factors such as patients' economic condition, trust in the platform and previous experience may be important predictors of premium payment, we specifically focus on service features that are visible on the platform. These transparent and visible features are one way that platform design and service delivery representation on the platform may potentially influence payment decisions. A better understanding of these visible service features on the platform should also help online medical service providers and platform administrators better understand and exploit their increasingly diverse service data. Accordingly, we ask the following three research questions:

(1) What are the observable features of online medical consultation services (specifically, characteristics of physicians and their interaction with patients which are visible on the platform) that are associated with patient payment, as opposed to free-trial-only appointments, in a freemium-based model?

(2) What is the relative importance of these features?

(3) How do these features interact, linearly or non-linearly, in relation to premium payment?

Our study contributes to online medical consultation by developing a holistic model with a system of service-related features in order to help the platform target high-value service providers (i.e., physicians in our context) and the type of services that attract premium payment. Effectively managing high-value service providers is essential for platform governance, and appropriate measures should be taken to identify them as early as possible and facilitate their retention on the platform (Voigt and Hinz 2016). In addition, since one of the social benefits of promoting digitally delivered medical services is increasing healthcare resource accessibility and equality, a better understanding online medical consultation services, especially those observable features of physicians and their interaction with patients, can help guide more efficient distribution and use of healthcare resources on the platform. In short, understanding the key service features of successful premium services can help the platform better manage its service focus and revenue generation mechanisms.

Methodologically, whereas previous research on online medical consultation almost exclusively used regression-based analysis, we adopt a machine learning approach to mine the massive consumer data. Although predictive modeling based on regression is useful to formalize statistical relationships between linearly arranged factors, the analysis relies on many data distribution assumptions, and only a limited number of independent variables can be included in the system to generate an isolated and linear prediction. Contrarily, we use a machine learning (ML) approach to investigate the usefulness of a bag of service features as a system of inputs (RQ1), and examine the relative importance of these features that may predict the outcome without too much constraint caused by data distribution or linear functions (RQ2 and RQ3). Since platform-based businesses are increasingly overwhelmed by a massive amount of fine-grained consumer behavior data with a high number of dimensions, a ML approach is especially useful in helping them to fully mine their service and consumer data to make better managerial decisions in a more holistic manner (Martens et al. 2016).

Our study will also contribute to the freemium literature. First, existing freemium literature on digital businesses almost exclusively focuses on *public* information goods such as mobile apps, software and music streaming. Public information goods are characterized by a substantial sunk cost to produce, yet almost negligible marginal cost for distribution (Shapiro and Varian 1998). However, *private* information goods that have non-negligible marginal cost have rarely been investigated. Online medical consultation is a type of private information good since the physicians need to provide highly personalized services with a certain level of quality for both free and paid versions, leading to relatively high marginal cost (Ozdemir 2007). Thus, previous literature on freemium may not fully cover the unique characteristics of digital healthcare delivery and the online consultation market, and our study will provide additional insights that complement those previous findings.

Second, existing freemium literature generally focuses on understanding freemium as a business model itself (e.g., cost and revenue structure), explaining consumers' willingness to pay, or improving freemium outcomes by designing better products or services (e.g., optimizing price or quality levels). Seldom did previous studies focus on service providers as a key component of the freemium model. Even for the mobile app market that frequently adopts a freemium strategy (a typical multisided market), previous research generally focuses on the product itself. However, in the context of online medical consultation, physicians as service providers are critical stakeholders in the digital ecosystem, as they execute the freemium services and generate revenue for the platform. When the services or products are provided on multisided platforms, service providers face fierce competition, as consumers can easily switch to other comparable service providers on the same platform. Thus, the multisided platform-based market is very different from single-sided digital businesses (e.g., proprietary software providers), and both the demand- and supply-side economies of scale are important. Our understanding of how service-related features influence the success of a freemium business is still limited. This study will provide initial evidence on how various service features are associated with premium payment (one important driver of physician and platform success) under a freemium business model.

In the following sections, we briefly introduce the characteristics of a freemium business model and the related literature on premium payment, especially in the context of online medical consultation. Then we describe our empirical setting and methodology for data analysis (e.g., machine learning procedures and algorithm selection). The results are presented in section 3.4, followed by discussion and conclusion.

3.2 Literature Review: Freemium Model and Premium Payment

Since this study is at the intersection of two streams of research, in the following section, we review online medical consultation literature to identify the variables that has been investigated to predict patients' service selection and payment, and freemium literature to understand the key mechanisms for explaining premium payment. Both literatures may be helpful to explain patients' payment in the context of online medical consultation on platforms using a freemium business model.

3.2.1 Patients' Selection and Payment in Online Medical Consultation

As has been discussed, online medical consultation may take various forms, ranging from online healthcare communities (which are usually free) to telemedicine (consulting with a designated physician through secured communication tools). Patients' acceptance of the service, satisfaction, and engagement are frequently examined topics in these contexts (e.g., Callahan et al. 1998; Dick et al. 1999; Jin et al. 2016; Maloney-Krichmar and Preece 2005; Misra et al. 2008; Müller et al. 2016; Song et al. 2015). These previous efforts may provide useful insights on patients' decisions regarding the selection and payment of a digitally delivered healthcare service. However, our focus is an emerging business model of online medical consultation on a multisided digital platform which is significantly different from Q&A communities or telemedicine, therefore requiring a closer examination of the influencing factors pertinent to the context.

Existing studies on patient's service selection and payment decision in online medical consultation are conducted mainly in a Chinese context, using the data from two major platforms (see Table 3.1). This is possible due to the popularity of such services in China and the accessibility of the data that is publicly available. Although both platforms employ a freemium strategy, only one study incorporated freemium elements in their research model, examining the impacts of free service offerings on physician revenue (Liu et al. 2018). Other studies generally treated free and paid service offerings (or appointments) as identical, or only look at the paid services. Patients' selection or payment outcomes are often measured by the total number of patient appointments or physicians' economic return.

Table 3.1 Studies on Patient Selection of Physicians or Payment in Online Medical Consultation

Study	Explanatory Variables	Outcomes	Theory	Method (platform)
Cao et al. (2017)	Service quality (# of existing patients); eWOM (Vote, star)	New patients' consulting intention	Elaboration likelihood perspective	Linear regression (Haodf)
Deng et al. (2019)	Physician online efforts; Physician reputation	Patients' choice of physicians	Social exchange theory	Linear regression (Haodf)
Guo et al. (2017)	Physician's status capital (title, hospital, city); Physician's decisional capital (# of services, frequency of service)	Physician's online economic return	Social exchange theory	Linear regression (Haodf)
Guo et al. (2018)	Doctor-patient interaction (strong tie, weak tie)	Physician's online economic return	Social capital and social ties	SEM (Haodf)
Li et al. (2019a)	Physician's online knowledge contribution; Physician's online reputation; Physician's offline status	Physician's online income	Signaling Theory	Linear regression (Haodf)
Li et al. (2016)	Physician's online reputation; Physician's online self-representation	Market concentration (sales & sales rank)	Information asymmetry theory	Linear regression (Haodf)
Li et al. (2019b)	Physician's online rating; Physician's online activeness; Physician seniority	Number of patients registered for a physician's service	Explanation drawn on service quality	Linear regression (Haodf)
Liu et al. (2018)	Amount of free services; Price; Physician's platform experience; Title	Physician's revenue	Elaboration Likelihood Model; Social Comparison Theory	Linear regression (Haodf)
Liu et al. (2016)	Physician's online reputation; Physician's offline reputation; Hospital's online & offline reputation	Amount of online service appointments	Signaling Theory	Linear regression (Guahao)
Lu and Wu (2016)	Online review (perceived treatment outcome & perceived physician attitude); Disease risk	Patients appointment totals	Explanation drawn on service quality	Linear regression (Guahao)
Wu and Deng (2019)	Leader capital; Medical team specification and credibility; Team collaboration	Medical team order quantity	Transactive memory perspective	Linear regression (Haodf)

Table 3.1 Studies on Patient Selection of Physicians or Payment in Online Medical Consultation				
Study	Explanatory Variables	Outcomes	Theory	Method (platform)
Yang et al. (2015)	Virtual gift & thank-you letter System-generated physician contribution value	Number of patients who consulted a physician	Signaling theory	Linear regression (Haodf)
Yang and Zhang (2019)	Virtual gift (as paid feedback) Thank-you letter (as free feedback)	Number of patients who consulted a physician	Signaling theory; Self-determination theory	Linear regression (Haodf)
Yang et al. (2019)	Previous consultation experience (response time, depth interaction, service content)	Patients' continuous consultation decision	Interpersonal trust development perspective	Logistic regression (Haodf)
Yu et al. (2017)	Physician credibility; Competition (location, group size); Endorsement (contribution score, popularity, profile)	Physician's online workload	N/A	Poisson regression (Haodf)

Two types of influencing factors are frequently examined – physician characteristics and patients' feedback. Since medical consultation is highly professional, credibility (or trustworthiness) is needed for physicians to attract patients and reduce their concerns for service quality. Thus, physician reputation – both online and offline – is the most frequently examined physician characteristic as the key antecedent of patient selection of physicians and payment (Liu et al.,2016). A physician's affiliation, seniority and location are usually used as proxies for reputation, which provide indicators of service quality and physician trustworthiness (Guo et al. 2017; Li et al. 2019a; Liu et al. 2018; Yu et al. 2017). Physicians' knowledge contribution (voluntary healthcare information contribution on the platform which is not relevant to the service) is another physician characteristic that has been examined to indicate a physician's self-representation online (Li et al. 2016; Li et al. 2019a; Yang et al. 2015).

Patient feedback, on the other hand, is the evaluation and e-word-of-mouth by previous patients on a physician's service quality or attitude, which is often measured by ratings, stars, reviews, and the amount of virtual gifts (Cao et al. 2017; Lu and Wu 2016; Yang and Zhang 2019). This kind of feedback mechanism is supported by the digital platform functionalities by which patients can express their opinions after the service and make them visible to other people. Previous patients' feedback may serve as endorsement signals and provides opportunities for future patients to learn about the service quality before the payment and actual service consumption.

Although less frequently examined, patient-physician interaction seems to be a promising influencing factor as well. As per social exchange theory, the theoretical lens that has been used by several previous studies, the depth of interaction and physician activeness (e.g., the amount of service or the frequency of service) are important capitals that indicate the ability and willingness of a physician to provide high-quality service (Guo et al. 2017; Guo et al. 2018; Li et al. 2019b; Yang et al. 2019).

In summary, existing research on online medical consultation focuses on limited types of factors that may contribute to patients' service selection and payment. These factors provide isolated explanations, so the business or service providers may have difficulties building a system of service-related features to predict payment. In addition, all of these studies follow a linear regression approach to investigate statistical relationships, which limits their ability to uncover complex dynamics between service features and outcomes. New methods are needed to complement existing research to explain and predict the designated outcome by considering service characteristics with a high number of dimensions.

3.2.2 Freemium Business Model and Premium Conversion

In this study, we aim to highlight the role of the freemium model in which patients have the option to stop further consultation after using up the free trials. Previous studies on freemium and related business models, such as sampling and versioning, provide useful insights regarding promoting payments and converting freebie consumers into premium subscribers. Herein, we review the key mechanisms and explanations that have been discussed in the relevant literature to inform the understanding of online medical consultation payment under a freemium business model.

Proposed by Wilson (2006), freemium has become an increasingly popular business model for online businesses to attract more users and generate revenue. Businesses that adopt a freemium model usually provide a free trial of the product or service (e.g., limited quantity, functionality, time of use, or add disturbances like advertisement), with the expectation of persuading consumers to pay for the advanced version with more features, unlimited use, or disturbance removal (Liu et al., 2014). Although free-trial consumers do not directly generate revenue, they can add value by attracting traffic and creating network externality (i.e., a product or service becomes more valuable when more users adopt it), thereby influencing subsequent consumer valuation and speeding up the product or service diffusion process (Gallaugher and Wang 1999; Jiang and Sarkar 2009). Thus, better leveraging the impacts made by free-trial consumers is critical to ensure sustainable revenue generation for a platform.

Bawa and Shoemaker (2004) summarized three types of direct impacts on follow-on sales that come from offering free trials. First, it might create an expansion effect, where offering free trials attracts new consumers who would not pay if there were no trials of the product. Second, an acceleration effect may occur where consumers pay for the product sooner than they would have if there were no free trials. Third, it might trigger a cannibalization effect where the offering of free trials may reduce the number of consumers who would have been willing to pay (e.g., unsatisfied trial experience, the fulfillment of the needs or no further service is needed). The first two are positive reinforcement, whereas the third has a negative influence. The net impact of offering free trials on subsequent premium purchases depends on the relative magnitude of these three mechanisms over time, and previous studies have reported that, in practice, a moderate conversion rate of 2% to 5% is considered sufficient to make up for the lack of profit from the non-paying consumers (Huang 2016; Kumar 2014). However, to the best of our knowledge, this has not been examined in the context of online medical consultation.

3.2.3 Freemium, Sampling and Versioning: How Free Trials Influence Premium Payment

Although there are limited previous studies on freemium-related issues in the context of online medical consultation, we can still draw implications from three closely relevant fields of research – sampling in marketing, versioning in IS and management and interdisciplinary research of freemium for digital businesses – to understand how offering free trials may influence patients' premium payment in online medical consultation. The comparisons between the three, as well as the common mechanisms that may influence payment, are presented in Table 3.2. These common mechanisms are a post-hoc categorization to show different types of explanation that were studied in the literature and are not meant to be an exhaustive or mutually exclusive list.

Sampling, versioning and freemium are similar in the sense that they are all multi-tiered pricing models with multiple segments of consumers (Huang 2016; Lyons et al. 2012). The close relationship between the three research areas is also indicated by the labeling of free parts in a freemium model as free samples or free versions. To differentiate, we use free trials versus premium for freemium research, free samples for sampling research and low-end version versus high-end version for versioning literature.

Table 3.2 Comparison between Sampling, Version and Freemium and Mechanisms Contributing to Payment			
	Sampling	Versioning	Freemium
Similarities	• Multi-tiered pricing • Multiple segments of consumers		
Research field	Marketing (perishable goods)	IS & Management (digital products)	Interdisciplinary (digital businesses in general)
General Purposes	Introducing and promoting new products	Consumer segmentation through vertically differentiated product quality	Scale up user base without expending costly resources on a traditional sales force.
Labels used in this study	Free samples vs. regular product	Low-end version vs. high-end version	Free trials vs. premium
Mechanisms Contributing to Payment with Study Examples			
Consumer awareness and learning [consumers pay attention to the products, remember their consumption experience and are motivated to consume the product again in future]	Heiman et al. (2001) Lammers (1991) Park et al. (1987)	Kannan and Kopalle (2001) Dey et al. (2013)	Liu et al. (2014) Nan et al. (2018)
Consumer valuation [consumers' cognitive evaluation that estimates the value or worth of the products]	Heiman et al. (2001) Lee-Wingate & Corfman (2010) Shampanier et al. (2007)	Bhargava & Choudhary (2008) Dey & Lahiri (2016)	Foubert &Gijsbrechts (2016) Wagner et al. (2013) Niemand et al. (2015) Nieman et al. (2019)
Product quality [consumers' perceived quality or the actual quality of the product]	Anselmsson et al. (2007) Sprott and Shimp (2004)	Shivendu & Zhang (2015) Raghunathan (2000)	Hamari et al. (2017) Nieman et al. (2019) Nan et al. (2018)
Consumer involvement [the degree to which a consumer is engaged with a product or service]	Park et al. (1987)		Oestreicher-Singer & Zalmanson (2013) Foubert &Gijsbrechts (2016) Datta et al. (2015)

As a marketing tool, offering free samples has been used a traditional strategy for new product introduction and promotion, with the expectation of facilitating immediate future purchase (Jain et al. 1995; Lammers 1991; Pawar et al. 2016). Versioning, on the other hand, is a popular strategy for promoting digital products such as software and mobile apps. The product or service under a versioning strategy is usually designed with vertically differentiated quality; thereby the business can segment consumers by considering preference heterogeneity, consumer self-selection and marginal costs (Bhargava and Choudhary 2008; Wei and Nault 2013; Wu and Chen 2008). Similarly, by allowing the coexistence of free and paid services or products, freemium enables the business to scale up consumer base without spending costly marketing effort (e.g., advertisement or marketing campaigns) on traditional sales forces (Kumar 2014).

We can observe several common and interrelated mechanisms in these three streams of research regarding how free products or services (or low-end versions) contribute to future sales. First, there may be a learning effect – offering freebies (or low-end versions) may increase awareness of the product in a highly competitive market and allow consumers to memorize their experience and create dependence on the product, which will further generate impacts on sales (Shapiro and Varian 1998). During the free sample consumption, consumers will learn from the experience and accumulate goodwill in the long run (Heiman et al. 2001). Thus, free sampling can be considered as a reinforcement vehicle that helps consumers make direct associations between the product and payment (Lammers 1991). The repetition of consuming free samples will help consumers establish a behavioral script to routinize payment (Park et al. 1987). Versioning literature has similar findings. Creating multiple versions of a product with different pricing tiers has impacts on consumer learning in that they may update expectations for future purchase by incorporating their experience of consuming the lower-end versions (Kannan and Kopalle 2001). Furthermore, when consumers' learning rate is high, a time-locked versioning strategy (e.g., free trials with a limited time) is considered to be optimal (Dey et al. 2013). Similarly, the free trials in a freemium model have been found to increase awareness of the product and provide opportunities for direct experience and learning (Liu et al. 2014; Nan et al. 2018).

In addition to the learning effect, offering freebies or low-end versions contributes to consumer valuation (i.e., consumers' cognitive evaluation that estimates the value or worth of the products or services), which may further facilitate sales. After consuming the free samples, the consumer may valuate the practical or utility value of the free product, form attitudes and perceptions, thus being more likely to pay in the future (Lee-Wingate and Corfman 2010; Wagner et al. 2013). Moreover, a free sample may bring consumers' attention to the positive cues related to consumption (e.g., good taste of food, the usefulness of a skincare product) and form goodwill after a positive consumption experience, thus increasing potential consumers' willingness to pay (Heiman et al. 2001). For versioning, although not all versioning strategy involves a free-tier product (e.g., a basic paid version to upgrade, low price individual edition vs. expansive enterprise edition), consumer valuation has also been found to play an important role. For example, Bhargava and Choudhary's (2008) study shows that for a monopoly firm that offers information goods with negligible variable costs, the choice of optimal versioning for business profitability depends on consumers' relative valuation of low and high-end versions. As consumers update their valuation and perceived fit after experiencing the low-end version, they may want more of the product capability and pay for the high-end version (Dey and Lahiri 2016).

The source for user valuation can be diverse. For example, in comparison with accepting a single paid version directly, consumers may perceive fewer sacrifices and more benefits when free trials are available (Niemand et al. 2015). Such a sacrifice-benefit evaluation may be derived from various mechanisms, such as the zero-price effect of the free-trials (i.e., consumers perceive free offerings as having irrationally higher value due to multiple psychological effects such as increased liking, motivation and sense of control, Shampanier et al. 2007), price-quality inference of the product (e.g., the expensive product should have higher quality, Niemand et al. 2019), or the functional fit between free and paid versions (i.e., the perceived similarity between free and paid services induces positive attitude toward the premium, Wagner et al. 2014). In addition, free trials may indirectly influence consumer valuation and payment through word-of-mouth. People who have tried the products may spread product-related reviews which influence other potential premium consumers' valuation of the product and adoption decisions (Foubert and Gijsbrechts 2016).

The third key mechanism is product quality, which may overlap with consumer valuation when consumers' perceived quality is the focus of study. Many consumers associate price with quality (Zeithaml 1988). Consequently, if consumers are not satisfied with the quality of the free-tier product (or low-end version) and thereby infer the quality of the paid or high-end version, they are less likely to pay due to this perceived lack of quality (Niemand et al. 2019). Previous studies in versioning have found that the driver for consumers adopting high-end versions is usually the limited utility of the low-end version, such as inconvenience and lack of required functionalities (Shivendu and Zhang 2015). Nan et al. (2018) have found that for the information goods market with a piracy concern, consumers' perception of the premium quality (versus their quality perception of the pirated version and the free version) influences the demand of premium. Not surprisingly, consumers' perceived quality of the premium service, such as reliability and responsiveness, is the key determining factor that contributes to retention and payment (Hamari et al. 2017). In addition to consumers' perceived quality, offering in-store free samples is generally considered as an opportunity in marketing to show the brand's quality, thus motivating purchase (Anselmsson et al. 2007; Sprott and Shimp 2004). Similarly, for products that are easily substituted by new competitors, multiple versions are helpful to clearly differentiate the uniqueness and quality (Raghunathan 2000).

The last mechanism, which is also often discussed in freemium literature, is consumer involvement (i.e., the degree to which a consumer is engaged with a product or service). Higher involvement with the product, even with free samples, has been suggested to be associated with various behavioral and cognitive factors, such as active information search and loyalty, which influences payment (Park et al. 1987). In general, consumers with more intensive involvement will become committed sooner and therefore will pay for the premium sooner, since they have invested more time and effort (Oestreicher-Singer and Zalmanson 2013). Involvement in the form of intensive free trial usage also facilitates consumer learning so that potential quality change between the product tiers is taken into consideration when making a payment decision, which reduces the sense of uncertainty (Foubert and Gijsbrechts 2016). More usage is also associated with consumers' value perception of the service or product. The previous study has found that although the average perceived value of free-trial consumers is lower than that of those who did not use the free option, free-trial consumers are more responsive to usage rates, such that the impacts of usage on retention is higher for those who experienced free trials (Datta et al. 2015).

In summary, when the consumers are able to test the quality or usefulness of the product freely, they will learn from their trial period and, if the trial experience is positive (e.g., they form positive attitudes, perceive high quality or value), will accumulate goodwill toward the product. Providing multi-tiered products, including free trials as the base version, can be a useful tool to exploit different groups of consumers' willingness to pay (Cox 2017). However, the effects of these approaches (including freemium) on promoting payment for information goods (e.g., online consultation service) can vary depending on when, where and under what condition it is offered (Chandukala et al. 2017),

Our focus on online medical consultation as a digital healthcare service is different from the existing multi-tier pricing literature in four ways. First, the sampling literature generally focuses on the in-store promotion of perishable goods that are usually restricted by time and location. However, online medical consultation is a type of information good with different distribution patterns (e.g., distributed online, without the time and location constraints, may differ in both quantity and quality), so offering free trials in such a context may yield different impacts on consumer experiences and perceptions.

Second, online medical consultation is a type of private information good, which is different from the frequently examined public information goods in the versioning literature. Public information goods such as software and mobile apps have substantial sunk cost to produce but negligible marginal cost for distribution, whereas online medical consultation as a private information good has non-negligible marginal cost due to its highly personalized nature (Ozdemir 2007). Because of such non-negligible costs for physicians, offering too many free trials may reduce their service quality of both free and paid consultations.

In addition, the quality differentiation assumption of versioning may not be fully applicable because the nature of healthcare service requires that physicians keep a certain level of quality (e.g., the ethical standards according to the Hippocratic Oath). In other words, providing a low-quality free software may not cause harm, but an incorrect free medical diagnosis and treatment plan can be harmful.

Lastly, many previous versioning studies follow a monopoly assumption where the market is characterized by one dominant provider (e.g., Bhargava and Choudhary, 2008), whereas we focus on a multi-sided platform on which consumers can easily switch to alternative service providers on the platform with comparable expertise and service quality (although a switch cost may exist). A freemium pricing structure on a multi-sided platform adds complexity to patients' decision making, and previous findings in versioning and sampling literature may not fully cover the unique characteristics of this emerging phenomenon.

In summary, a patient may experience reluctance due to various factors such as information asymmetry, lack of domain knowledge and difficulties in evaluating service quality (Abu-Salim et al. 2017; El-Manstrly 2016; Rose and Samouel 2009). Providing initial free services can be a useful strategy to attract those who are hesitant to adopt premium services online. If patients can form a positive attitude through trials (i.e., service valuation), they are more likely to pay for the premium (Huang & Korfiatis 2015). This positive attitude formation may be influenced by the awareness of physicians, patients' involvement during the service, and service quality as indicated by physicians' activities and their online and offline reputation (e.g., physician response, previous offline service).

3.3 Methodology

To examine the key service features that may contribute to premium payment, we use a machine learning approach to explore the associations and predictive power of various consultation service-related features regarding service awareness, patients' involvement, service valuation and service quality. Whereas previous studies in similar contexts generally used a linear regression-based approach, the ML approach is particularly useful in our context due to its ability to mine massive fine-grained behavior data (Martens et al. 2016), and it allows the embedding of previous insights (i.e., literature-based feature selection and model building) into versatile data mining processes and cross-validation of the predictive model (Kitchens et al. 2018). The various classifiers following different ML philosophies for a binary classification problem provide complementary views of how the model with our selected features can help us understand premium payment.

3.3.1 *Empirical Setting and Data Collection*

Our empirical setting is a multisided online medical consultation platform based in China. It is one of the largest medical platforms and has attracted more than half a million physicians from over 9,400 hospitals to set up their profiles and provide consultation services on the platform. As of 2018, more than 50 million patients uploaded healthcare records to seek medical consultation services. The platform follows a freemium model under which a patient can gain up to three sessions of free consultations (i.e., three free trials) with one particular physician through two channels: either have previous offline hospital visit experience with the same physician (i.e., returning patients) or allow the platform to assign a physician when the patient initiates the consultation (i.e., assigned patients). Patients have to pay for the premium under two conditions: (1) he/she used up the three free trials with a particular physician and wants additional service from that physician; (2) he/she selects a physician when initiating the consultation without any previous offline visits. In this latter case, the patients may or may not have experienced free services from other physicians. Switch costs exist when a patient changes a physician in order to get more free consultations, which includes extra time and effort to report healthcare history, describe symptoms, establish a relationship with the new physician, and learn a new collaboration style. The free service is based on asynchronous picture-and-text communication, which allows a physician to significantly delay giving a response when busy. Premium packages are diverse and are supported by more synchronous communication such as instant messages and phone calls with various time combinations (e.g., unlimited messages within 48 hours, monthly subscription, 15-minute phone call).

We collected consultation records between patient-physician pairs that cover a period of twelve years (from when the platform launched in August 2006 to August 2018). To keep the data collection and cleaning to a manageable size, we only collected data from the three departments that have received the most visits[14] (i.e., Pediatrics, Gynecology, Dermatology) across six geographic areas – three areas with the richest healthcare

[14] According to the Healthpoint 2017 Industry Report: http://www.healthpoint.cn/article_detail/57925

resources (Beijing municipality, Guangdong province, Zhejiang province) and three remote areas with the fewest healthcare resources (Shanxi province, Tibet province, Qinghai province)[15]. For these departments and areas on the platform, all patient-physician consultation records during this time period were collected, resulting in more than 6.8 million records from 7,727 unique physicians[16]. The number of consultation records for our dataset by month is shown in Figure 3.1. We observed that the amount of consultation service offerings significantly increased after 2009, possibly due to the maturity of the platform or industry. To avoid systematic differences, we excluded the records before January 1, 2009 (N=17,397).

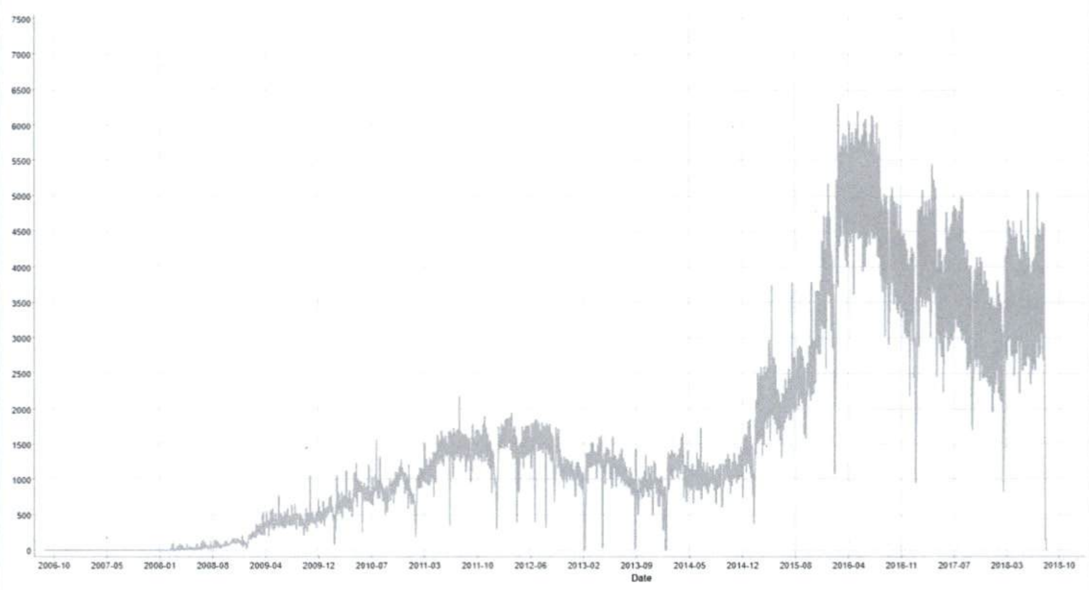


Figure 3.1 Number of Consultation Records by Month

In our data, each record is a consultation history which includes picture-and-text based dialogues and service purchase records between patient i and physician j (see an example screenshot in Figure 3.2). As mentioned, certain patients have free trials on the platform. One free trial includes one conversation between the patient and the physician. One

[15] These areas are selected according to the healthcare resource measures in the 2017 Yearbook of Health (i.e., the number of IIIA hospitals and the number of certified physicians per province or municipality). These areas cover approximately 16.62% of the whole population in China.

[16] There is no unique identifier for patients available. Thus, we do not know the total number of unique patients in our dataset, but each record represents the whole consultation history of a unique patient-physician pair.

complete conversation is one turn by the patient (which can include multiple posts) followed by one turn by the physician (which can also include multiple posts). At the time of our data collection, the website provides visible tags to indicate whether the record includes payment (see Figure 3.3). In addition to the basic freemium rules mentioned above, a physician may give a patient additional free consultation (in most cases, one or two free trials), but the platform does not encourage this practice and the physician needs to apply so that the platform gives the additional free services. Also, some premium services do not provide a tag (e.g., family doctor package) and some are set as "private", thereby no payment tag is provided. Thus, to avoid categorizing a record under the wrong type (i.e., free-trial vs. paid), any record that do not include a clear payment tag on the platform is removed from our dataset. We only include the following four types of tags which include most types of short-term paid services – 48-hour unlimited text & picture consultation, one-time phone-based consultation, one-question-and-one-answer, and check-in[17]. Each record can have more than one tag. The first three tags indicate payment, and the "check-in" tag indicates the free services when a returning patient had a previous offline consultation with the physician. Given these freemium rules and this tag system, Table 3.3 presents different types of patient-physician relationships.

[17] Other types of tags include private (for particular dialogues), thank-you letter, and virtual gift.

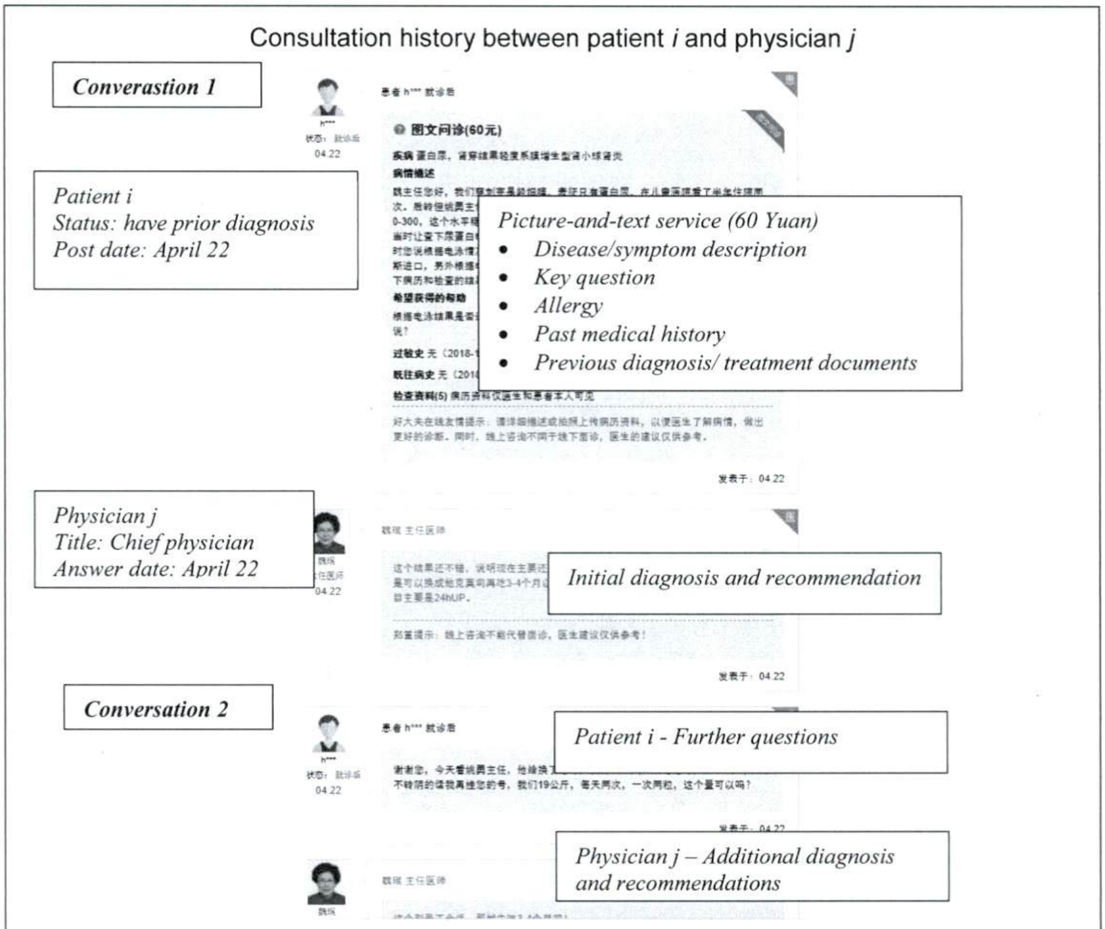


Figure 3.2 A Screenshot of Consultation History between Patient i and Physician j

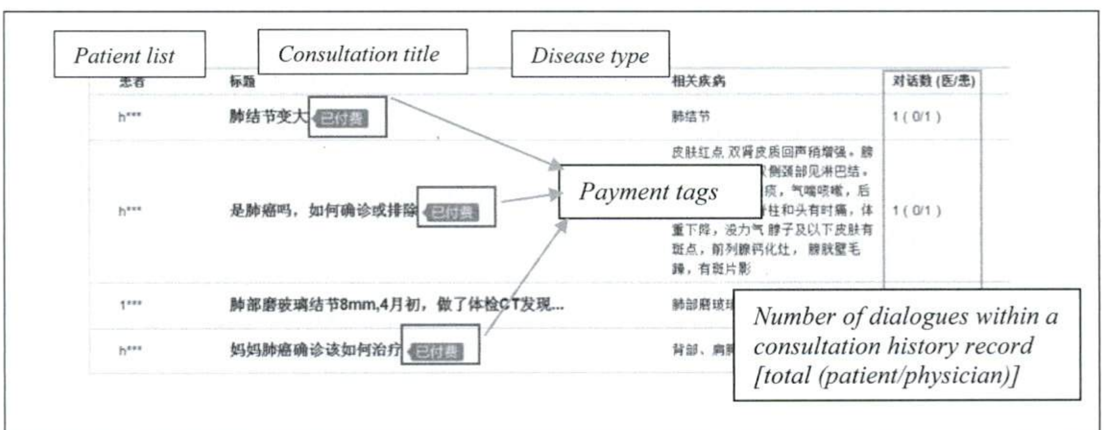


Figure 3.3 Screenshot of Transaction Records and Labels

Table 3.3 Types of Patients and Services		
	Free-trial-only	With premium subscription
Have an offline relationship	(1) Returning patients All offline return patients can enjoy three free trials with the same physician who gave them offline services. These patients usually use this opportunity to clarify the unsolved issues during their offline hospital visit or for follow-ups **Identification**: With a "check-in" tag/ No "payment" tag	(2) Returning patients who pay for the online premium service If a returning patient wants to continue the service with a particular physician after using up the free trials, he/she has to pay for the premium. **Identification:** With a "check-in" tag/ With at least one "payment" tag
No offline relationship	(3) Platform-assigned patients If a patient does not select a physician (i.e., the physician is assigned by the platform) when he/she initiates the consultation, the patient enjoys three initial free trials. These patients may have tried the free service from other physicians and switched to the current one or may be new to the platform. **Identification:** No "check-in" tag/ No "payment" tag/ Turns of conversation less than four*	(4) Online patients who pay for the premium service Two types of patients exist: 1. The patient who did not select the physician and used up the free trials must pay if he/she wants to obtain further services from that physician. 2. A patient who selects the physician has to pay (no free trials) when he/she initiates the consultation. **Identification**: No "check-in" tag/ With at least one "payment" tag

* Some physicians may decide to give extra free consultations (beyond three) to certain patients. These extra free consultations are not indicated by a tag on the system. To ensure that these special cases (which cannot be consistently identified) are removed from our data, all type 3 services (Platform-assigned patients) who had four or more free consultations were removed from the dataset.

3.3.2 Key Predictive Features

Inspired by previous research on sampling, versioning and freemium, we extracted consultation record features that are potentially related to the interrelated mechanisms highlighted in the literature review section. These features are all displayed on the platform and are visible by platform users (e.g., visitors, patients, physicians), thus allowing potential patient learning and valuation prior to the actual consumption of the consultation service. Since the number of physicians available on the platform is very high, we expect service awareness to be related to physician online and offline reputation – the more prestigious the physician, the more likely patients will be aware of them. We expect physicians' affiliations (e.g., from top hospitals or remote "unknown" hospitals) and titles (e.g., senior chief physician vs. junior associates) to serve as proxies for offline reputation. We extracted a physician's affiliated hospital ranking, hospital location, department and seniority (as represented by the title). For online reputation, we extracted a physician's online tenure and service intensity. These features may be also associated with service quality, since they reflect a physician's credibility or trustworthiness. We expect that those from good hospitals and longer tenure on the platform with more patients per month will be perceived as more credible. The definitions and coding are presented in Table 3.4.

Table 3.4 Key Predictive Features and Coding Description	
Feature	Description
Physician-related	
Hospital ranking	**[Ranking1]** Equals 1 if IA or IB hospital, 0 otherwise **[Ranking 2]** Equals 1 if IIA or IIB hospital, 0 otherwise
Physician seniority	**[Title1]** Equals 1 if chief physician, 0 otherwise **[Title2]** Equals 1 if associate chief physician, 0 otherwise
Hospital location	**[Loc]** Equals 1 if healthcare resource-rich areas, 0 otherwise
Physician tenure	**[Tenure]** The number of months the physician has been registered on the platform
Service intensity	**[Intensity]** The average number of patients served per month during the physician's tenure (=total patients served / Tenure)
Patient-related	
Previous formal examination	(a function provided by the platform allowing patients to reveal their medical status) Status 1: the patient reveals that he/she has had no formal healthcare examination before the online consultation. Status 2: the patient reveals that he/she has had a formal healthcare examination before the online consultation. Status 3: the patient chooses to set their previous examination history as private (i.e., the detailed consultation information is hidden and not directly visible by other patients) (coded into dummies) **[PriorExam]** Equals 1 if none, 0 otherwise **[Private]** Equals 1 if set as private, 0 otherwise
Offline connection	**[Offline]** Equals 1 if the patient used the "check-in" function when he/she initiated the consultation, which indicates a previous offline connection with the physician.
Service-related	
Service duration	**[Duration]** Number of days between the initial post and last post of patient *i*'s interaction with physician *j*
Total dialogue	**[TotalD]** Total number of posts within patient *i*'s interaction with physician *j*
Patient posts	**[PatientP]** Number of posts initiated by patient *i* within that patient's interaction with physician *j*
Physician posts	**[PhysicianP]** Number of posts initiated by physician *j* within patient *i*'s interaction with physician *j*
Response rate	**[Response]** the rate of response of a physician (= PhysicianP/ TotalP)
Question frequency	**[Question_frq]** The average number of posts by the patient per day during patient *i*'s interaction with physician *j* (=PatientP/ Duration)

Table 3.4 Key Predictive Features and Coding Description	
Feature	Description
Answer frequency	**[Answer_frq]** The average number of answers (including notifications and reminders) by the physician per day during patient *i*'s interaction with physician *j* (=PhysicianP/Duration)
Social return	**[Social]** Equals 1 if patient *i* gave any virtual gift to physician *j* at any time during patient *i*'s interaction with physician *j*

Note: The labels in square brackets are the variable names.

For service quality and patient involvement, we also consider online and offline patient-provider interactions. For some patients, online consultation is an extension of their offline hospital visits, so taking into account the offline relationships between patients and providers may help explain why some groups of patients pay while others do not. We consider whether a patient has previously had a formal medical examination for their current issues (not necessarily with a physician on the platform). We also consider whether a patient has had a previous offline examination with this physician from the platform. The former feature is useful in that it indicates whether official healthcare records are available. Therefore, the patient knows more precisely the service he/she is seeking. The offline connection, on the other hand, indicates whether a relationship has been already developed prior to the online consultation.

The abovementioned features are defined before the online consultation and can therefore serve as predictors of premium payment, whereas online interaction develops during the consultation. Although reverse causality issues may exist (thus they cannot be claimed as predicting factors), the association between online service quality-related characteristics and premium payment can still inform physicians and the platform on how to improve their service attractiveness. We consider both patient involvement (i.e., number of patient posts and question frequency) and physician response (i.e., number of physician posts and answer frequency). The associated service duration and the total dialogue are also extracted since they reflect service efficiency. Social return (i.e., virtual rewards for physicians such as a thank-you letter and virtual gifts) is also extracted as another aspect of involvement since it reflects non-monetary commitment, which may develop into monetary commitment (i.e., premium payment).

3.3.3 Outcome and Machine Learning (ML) Tasks

Our focal outcome is whether or not a consultation includes premium payment. A consultation record may include multiple premium payments, but at the current stage, we do not consider intensity or type of payment. Accordingly, the objective of our machine learning task is to solve a binary classification problem – classifying consultation history records into free trials only (labeled as *free*) or those including some type of premium payment (labeled as *paid*).

3.3.4 Analysis Steps

The key analysis steps are presented in Figure 3.4 and Table 3.5 below. After data extraction based on the features outlined above, initial cleaning and screening procedures are conducted to ensure the quality of data input. This includes transforming the data (i.e., string data, non-English characters, non-compatibly formatted data) into the right representations (e.g., extracting numeric values, turn strings into appropriate categorical data, dummy variable coding, new variable calculation). Records with too many missing values are deleted (since our sample size is large enough to be representative), and outliers (mainly due to errors in the scraping process) are removed using 95% quantile as the threshold.

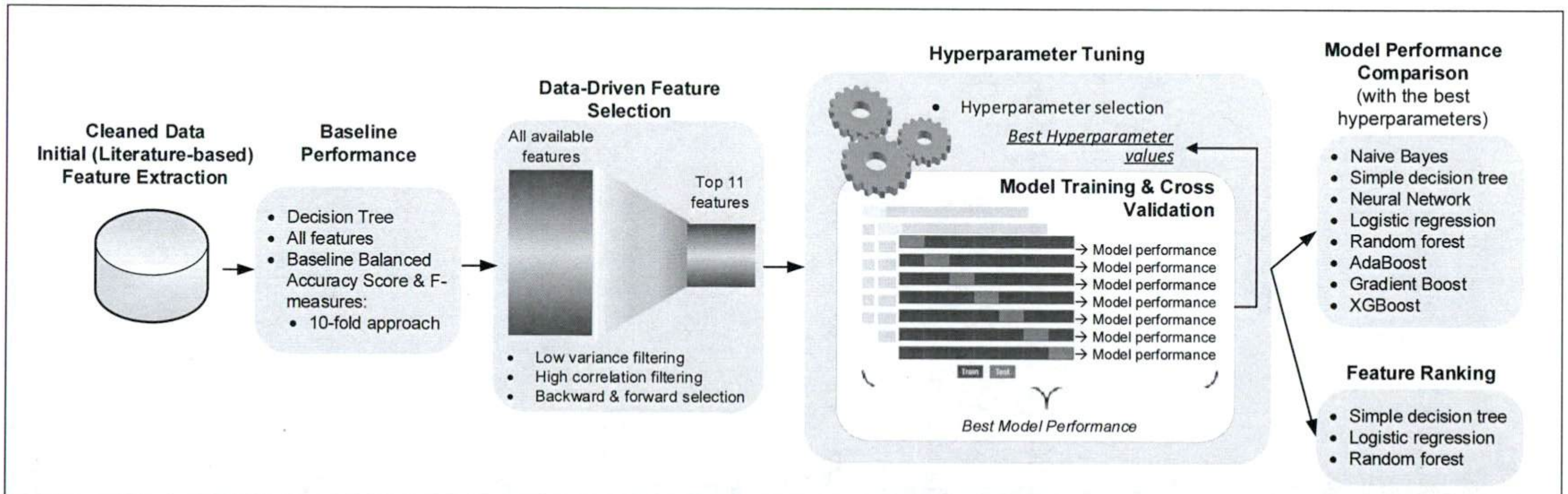


Figure 3.4 Analysis Pipeline

Table 3.5 Explanation of Methodological Steps	
Methodological Steps	Purposes
Data extraction (Literature-based feature selection)	Extract appropriate variables from the scraped dataset based on the key predictive features discussed above.
Data cleaning and screening	Transform the raw data into appropriate data representations to ensure the quality of inputs and outcome. This process includes: - Extract numeric value from string format data - Transform categorical data into dummies - Remove records with missing values - Detect and remove outliers - Prepare the outcome variable: turn string format tags into a categorical variable
Data-driven feature selection	Dimension reduction: retain the most useful information to generate a parsimonious prediction model that can achieve good classification performance. Decision Tree algorithm is used. Dimension reduction techniques include: - Low variance filtering - High correlation filtering - Backward feature elimination - Forward feature construction
Model training and validation	Execute the training and validation tasks in loops to optimize the hyperparameters, find the model parameters and compare the performance of different ML algorithms. The following steps are executed iteratively in loops.
Classification algorithm selection	A variety of common algorithms are used to ensure the accessibility of our approach. We applied Naïve Bayes, Decision Tree, MLP Neural Network, Logistic Regression, Random Forest, AdaBoost, Gradient Boosting and XGBoost to represent different machine learning philosophies for classification problems.
Hyperparameter tuning & model training	Determine the optimal setting of the ML algorithm in order to achieve better performance during model training and prediction. To account for the overfitting issue, we used a 10-fold cross-validation approach (i.e., iteratively train the selected nine sub-sets and then evaluate on the 10^{th}).
Model validation	Test the model (the feature configurations) by predicting the payment outcomes in the validation set.
Model evaluation	The performance of the predictive models with the best hyperparameters is evaluated based on commonly accepted metrics: recall, specificity, precision, F-measure, accuracy, balanced accuracy, AUC of ROC and Cohen's kappa.

Next, identifying the right features to use is the key challenge for effective ML classification – it is an essential step in retaining useful features and thus improves computation efficiency without over sacrificing classification task performance. Although we expect all the features listed in Table 3.4 to be useful, some of them may have low explanatory value and high multicollinearity, rendering them less useful. To make the model more parsimonious, we follow Silipo et al.'s (2015) approach[18] and run four tests – low variance filtering, high correlation filtering, backward feature selection, and forward feature selection – to eliminate less useful features before formal model building. The first two are *filter* approaches in which the features are evaluated based on data characteristics (i.e., variance and correlation), and the last two are *wrapper* approaches, which use ML algorithms and statistical re-sampling techniques to decide the importance of the features (Yu and Liu 2003). The objective of this procedure is to find suitable features that are highly correlated with the outcome but ideally uncorrelated with each other (Hall 2000). We use a decision tree classifier, which is a classic ML algorithm for feature selection (Dash and Liu 1997). Since the objective at this stage is not to assess or improve classification performance but to find a baseline for model training and selection, we do not perform hyperparameter tuning at this stage.

Model training and validation is the core of our analysis. This includes three nested tasks: hyperparameter optimization, model training and testing. In addition to the three basic algorithms used in feature selection, six additional algorithms that are suitable for classification problem are used (see Table 3.6). These algorithms differ in their classification approaches and required data distribution. For example, Naïve Bayes is a generative approach that learns the joint probability distribution p(x,y) and then uses Bayes' rule to calculate p(y|x), whereas the others are discriminative approaches that learn the conditional probability distribution p(y|x) directly. MLP neural network and Logistic Regression are parametric approaches that make prior assumptions on data distribution and the form of mapping functions (e.g., a linear function with a fixed number of parameters), whereas the others are non-parametric, so the complexity of the model may grow as more training data is input into the system. Comparing to these algorithms, the random forest and boosted trees are ensemble learning approaches (i.e., using multiple classifiers in the system). However, they differ in learning approaches (i.e., bagging vs. boosting), and AdaBoosting and Gradient Boosting (also XGBoost) differ in boosting

[18] Silipo et al. (2015) suggested seven techniques, and we did not use the missing value, PCA and random forest-tree approaches because (1) we have already eliminated records with missing value before the analysis, (2) PCA transforms the data into orthogonal components which are feature combinations that are difficult to interpret, and (3) random forest is one of the main algorithms for full model analysis.

approaches (i.e., focus on data distribution vs. residuals). These differences reflect different ML logic and decision boundary detection approaches (Murphy 2012).

Table 3.6 Machine Learning Algorithm Comparison				
ML Algorithm	**Classifier System**	**Classification Approach**	**Parametric**	**Explanation**
Naïve Bayes (NB)	Single	Generative	Yes	Use Bayes' theorem to model the distribution of input features and calculate the posterior probability of the hypothesis. The features are assumed to be independent.
Decision tree (DT)	Single	Discriminative	No	Find the decision boundary by bisecting the space into hyper-rectangles.
MLP neural network (NN)	Single	Discriminative	Yes	Mapping the input features to output classes by composing simpler data representations (i.e., creating hidden layers that extract abstract information from the visible input layer).
Logistic regression (LR)	Single	Discriminative	Yes	Find a single line decision boundary (linear or not linear) using Sigmoid function.
Random forest (RF)	Multiple (bagging)	Discriminative	No	Decision tree + Bagging: Trees are grown in parallel by randomly selecting samples (bootstrapping) with a subset of features that can give the best split on the training data. Predictions are aggregated from these sub-models.
AdaBoost tree (ADA)	Multiple (boosting)	Discriminative	No	Decision tree + Boosting: Bootstrap data and grow trees sequentially by iteratively updating data distribution with upweighted previously misclassified data (i.e., previously misclassified observations are more likely to appear in the next bootstrap)
Gradient Boost tree (GB)	Multiple (boosting)	Discriminative	No	Decision tree + Boosting: Instead of updating the data distribution to highlight the misclassified observations, gradient boosting approach focuses on the residuals and uses a gradient descent method to minimize the loss function.
XGBoost tree (XG)	Multiple (boosting)	Discriminative	No	Gradient boosting technique being optimized for speed and performance (i.e., apply penalties when updating trees and residuals).

These ML algorithms are classic and common approaches that are available in multiple data analytics platforms and packages (e.g., R, Python libraries, SAS, RapidMiner). This ensures the accessibility of our approach and prediction model to general data consumers

(Kumar et al. 2018; Lukyanenko et al. 2019). Using multiple ML algorithms also help to cross-validate the model performance and to provide implications on feature importance and feature configurations.

Depending on the ML algorithm, different sets of hyperparameters need to be configured to ensure that the algorithm reaches its best classification performance. Hyperparameters are external to the model and need to be decided upon first so that they can be used later during model parameter estimation (i.e., model parameters are learned from data, whereas hyperparameters are tuned by researchers). Searching for the optimal hyperparameter value is a process of trial-and-error, depending on the dataset (although sometimes rules-of-thumb may be applied). The hyperparameter optimization process (or called model tuning) is integrated with model training and validation in which we use a k-fold cross-validation approach[19] (we set k=10) to partition and train the subsets of the sample. We then compare all of the models and use the best hyperparameters for training and validating the full data set. Compared to traditional percentage-based one-time partitioning (e.g., 80%-20% partition), k-fold cross-validation is suggested to be a better approach to avoid overfitting because the classifier under a one-time partitioning is more likely to be tuned to fit particular sample characteristics (Cooil et al. 1987; Cui et al. 2008).

The results of the classification are evaluated with commonly accepted performance metrics (Abbasi and Chen 2008; Batista et al. 2004; Kumar et al. 2018; Larsen and Bong 2016; Sokolova and Lapalme 2009). In addition to traditional measures such as precision, recall, F-measure, and area under the ROC curve (AUC), we also report balanced accuracy and Cohen's kappa since they are considered to be better performance measures for imbalanced data (Brodersen et al. 2010; García et al. 2009; Jeni et al. 2013; Jiao and Du 2016; Powers 2011). An explanation of each measure is presented in Table 3.7.

---

[19] K-fold cross-validation provides a low-bias model selection solution by randomly partitioning the sample into k equal-size disjoint subsets, using (k-1) subsets as the training data and then making a prediction on the remaining subset. The process is repeated k times; thus, all data are used for both training and testing without using replacement, which is different from other sample construction selection approaches such as percentage-based partitioning and bootstrapping.

Table 3.7 Evaluation Measures and Explanation		
Measure	Calculation	Explanation
Precision	TP / (TP + FP)	The positive predictive value (also called purity). It shows the percentage of samples correctly labeled as positive (i.e., *paid*) out of all predicted positive.
Recall	TP / (TP + FN)	The true positive rate (also called sensitivity). It shows the percentage of samples correctly labeled as positive out of all true positives.
Specificity	TN / (TN + FP)	The true negative rate (also called selectivity). It shows the percentage of samples correctly labeled as negative out of all true negatives (i.e., *free*).
F-measure	$\frac{2\text{* Precision * Recall}}{\text{Precision + Recall}}$	The harmonic mean of precision and recall. It is useful for unbalanced data since it takes into account the cost of both false negatives and false positives.
Balanced accuracy	(Recall + Specificity)/2	The average of recall and specificity. It combines correctly predicted positives and correctly predicted negatives.
Cohen's kappa	$\frac{\text{Observed accuracy – expected accuracy}}{\text{1- Expected accuracy}}$	It measures the agreement between algorithm-predicted classification and true classification, ranging from -1 (completely disagree) through 0 (random) to 1 (perfect agreement).
Area under the ROC curve (AUC)	$P(X_+ > X_-)$ X+: the score of a random positive sample X-: the score of a random negative sample	AUC represents the probability that the ML model ranks a random true positive sample more highly than a random true negative sample (i.e., true positive rate vs. false positive rate). The closer to 1, the better the prediction. AUC=1 means 100% true positive and 0% false positive.

Note. TP – true positive; FP – false positive; FN – false negative; TN – true negative.

3.4 Results

3.4.1 Descriptive Characteristics of the Study Sample

After excluding the samples with unqualified tags (N=3,843,049), too many missing values (N=84,582) and outliers (N=674,767), 1,582,564 qualified records remain for further analysis. Among these records, 1,089,662 are free-trial-only, whereas 492,902 involve at least one premium payment. When performing ML tasks, an unbalanced dataset may be considered a limitation in that the minor class may not be learned adequately by the algorithm since it will get more experience from the majority class. However, since our sample size is large enough, and the unbalance rate is not particularly high (approximately 1:2.2), the small proportion of records with payment should not cause severe problems. In addition, remedies such as boosting, bagging and repeated resampling techniques are implemented in several ML classifiers to cross-validate the results. Performance measures that are less sensitive to unbalanced data are also used in addition to the traditional measures (He and Garcia 2008)[20]. Table 3.8 presents the summary statistics of the preprocessed data and features.

Table 3.8 Summary Statistics of the Observable Features

	Mean (SD)				Min – Max		
	All	Free-only	Paid	p-value*	All	Free-only	Paid
Physician-related							
Hospital ranking 1	0.947 (0.224)	0.943 (0.232)	0.955 (0.207)	1.336	0 – 1	0 – 1	0 – 1
Hospital ranking 2	0.022 (0.145)	0.025 (0.156)	0.014 (0.119)	<0.001	0 – 1	0 – 1	0 – 1
Physician title 1	0.465 (0.499)	0.434 (0.496)	0.533 (0.499)	<0.001	0 – 1	0 – 1	0 – 1
Physician title 2	0.326 (0.469)	0.342 (0.474)	0.291 (0.454)	<0.001	0 – 1	0 – 1	0 – 1
Hospital location	0.937 (0.243)	0.925 (0.263)	0.962 (0.191)	<0.001	0 – 1	0 – 1	0 – 1
Physician tenure (month)	74.724 (35.836)	71.984 (36.111)	80.782 (34.456)	<0.001	0 – 125	0 – 125	0 – 125
Service intensity	54.089 (48.951)	52.423 (47.949)	57.772 (50.904)	<0.001	0.008 – 207.062	0.008 – 207.062	0.009 – 207.062
Patient-related							
PriorExam	0.187 (0.39)	0.06 (0.238)	0.468 (0.499)	<0.001	0 – 1	0 – 1	0 – 1
Private	0.077 (0.267)	0.066 (0.249)	0.101 (0.301)	<0.001	0 – 1	0 – 1	0 – 1
Offline connection	0.714 (0.452)	0.873 (0.332)	0.363 (0.481)	<0.001	0 – 1	0 – 1	0 – 1
Service-related							
Service duration (day)	66.749 (204.027)	55.374 (190.328)	91.898 (229.458)	<0.001	1 – 4,297	1 – 4,297	1 – 3,824
Total dialogue	7.444 (6.379)	6.268 (5.026)	10.044 (8.061)	<0.001	1 – 35	1 – 31	1 – 35
Patient posts	5.792 (5.104)	5.052 (4.384)	7.429 (6.104)	<0.001	1 – 28	1 – 28	1 – 28
Physician posts	1.651 (1.776)	1.215 (1.111)	2.616 (2.46)	<0.001	0 – 7	0 – 3	0 – 7
Response rate	0.187 (0.163)	0.179 (0.158)	0.203 (0.173)	<0.001	0 – 0.875	0 – 0.75	0 – 0.875
Question frequency	1.195 (1.428)	1.196 (1.238)	0.923 (1.071)	<0.001	0 – 5.5	0 – 5.5	0 – 5.5
Answer frequency	0.222 (0.331)	0.24 (0.345)	0.185 (0.294)	<0.001	0 – 1.25	0 – 1	0 – 1.25
Social return	0.179 (0.383)	0.179 (0.383)	0.178 (0.383)	0.263	0 – 1	0 – 1	0 – 1

*Mean difference between *free* and *paid*. Levene's test indicates that the variances are not equal between free-trial-only and paid groups across all variables.

[20] Although we expected that unbalanced data would not influence the results, we perform four additional analyses with balanced data to verify the robustness of our results (presented in section 3.4.6).

3.4.2 *Feature Selection*

Before performing the feature selection procedure, we train and validate the data with decision tree algorithm using a 10-fold cross-validation approach with all the features. The purpose is to set the baseline level for classification performance so that we can compare the performance of various dimension reduction approaches and balance the feature retention options. The decision tree algorithm reaches a balanced accuracy of 95.9% (F-measure=0.965, Recall= 0.959, Precision= 0.971).

Then we run the models iteratively with four feature selection techniques. First, dimensions with very low variance (i.e., almost a constant value) can be eliminated since they contain little information to discriminate the classes. To find the optimal threshold for elimination, we apply the decision tree algorithm in loops[21] with 10-fold cross-validation, and the threshold that can maximize the balanced accuracy score is selected for deciding the feature retention. As shown in Table 3.9, a variance of 2% is considered as a low-bound threshold, and if dimension reduction is executed based on this threshold, after eliminating the low variance features, the classification (balanced) accuracy can reach 97.85%.

Second, highly correlated dimensions can be eliminated since they contain similar information. Thus, removing one of the two highly correlated dimensions can simplify the input space without sacrificing the predictive power for the future. Pairs of correlations are calculated after variable normalization since the calculation of the correlation coefficient largely depends on the data range, and one of the two highly correlated dimensions is eliminated if it goes above the correlation threshold. Similar to the identification of low variance threshold, the optimal correlation threshold is decided by running a loop of training and validation tests with three algorithms to achieve the highest classification accuracy. The correlation matrix is presented in Appendix C - Table C.1. The optimal threshold for dimension reduction decided by the algorithm is 0.3. Five

[21] First, the variance of each dimension is calculated based on normalized data. Then, low variance dimensions are filtered out in the subsequent training and validation tasks using a Decision Tree algorithm. Such screening process is running in a loop by setting the threshold values from 0.01 to 0.1 with incremental increase of 0.01 every time when a training and validation algorithm is applied.

features are retained (see Table 3.9) based on this threshold, which can yield a balanced accuracy level of 78.1% in the classification task on the original data based on the Decision Tree algorithm. However, the algorithm-decided feature retention may be too strict, resulting in too much information lost (i.e., only the most parsimonious solution is given). Thus, we try a manual procedure in order to retain more feature information. We manually select features with low correlations (below 0.3) based on the correlation table and input them all together into the training model. Ten features were used which can yield a balanced accuracy of 97.86%.

Third, backward feature elimination removes one input feature at one time during the training and validation loop, and each feature removal should minimize the increase of error rate in the classification performance. When setting the expected accuracy threshold for feature retaining at 96% (the baseline balanced accuracy for decision tree algorithms), five features are retained as the most parsimonious solution[22] that yields a 97.8% balanced accuracy.

Fourth, contrary to the backward approach, forward feature selection starts from inputting one feature and adds one new feature at a time in the iterations to obtain the best classification accuracy. Similar to the backward approach, when we set the target accuracy threshold at 96%, four features are selected as the most parsimonious solution. The prediction on the validation set yields a 97.8% balanced accuracy level as well.

The four approaches listed above did not give us consistent feature selection results, which is within our expectation because they follow different philosophies to reduce the dimension. Since forward and backward approaches gave us the most parsimonious solutions – adding peripheral features will not reduce the model's prediction performance – we decide to retain additional features that are selected by both low variance filtering and one of the high correlation filtering approaches. Consequently, in response to our first research question regarding which observable features of online medical consultation services are associated with patient payment, our feature selection analysis suggests that

[22] Multiple combinations of the features can yield similar expected prediction accuracy, and we only report the simplest combination here.

the key observable features are: *ranking 2, title 1, PriorExam, private, offline connection, total dialogue, patient posts, response rate, question frequency, answer frequency,* and *social return.* These features are used as the inputs in the analyses which follow.

Table 3.9 Feature Selection Results

	Low Variance filtering	High correlation filtering	High correlation filtering (manual)	Backward feature elimination	Forward feature elimination
Hospital ranking 1	√				
Hospital ranking 2	√	√	√		
Physician title 1	√	√	√		
Physician title 2	√				
Hospital location	√				
Physician tenure (month)	√				
Service intensity	√				
PriorExam	√		√		
Private	√	√	√	√	√
Offline connection	√		√	√	√
Service duration (day)			√		
Total dialogue	√	√			
Patient posts	√		√	√	√
Physician posts	√				
Response rate	√		√	√	
Question frequency	√		√		
Answer frequency	√	√			
Social return	√		√	√	√
Balanced accuracy (Decision tree with 10-fold cross validation)	97.85%	78.13%	97.86%	97.8%	97.8%
Threshold	0.02 (loops)	0.3 (loops)	0.3 (once)	-	-

3.4.3 Hyperparameter Tuning

The optimal value of hyperparameters is decided through a cross-validation process. This involves several subjective decisions and trial-and-error. For example, hyperparameters for a decision tree may involve the maximum depth of a tree, the minimum sample size of a node, the minimum sample size of a leaf, and the maximum number of features included for a split (Shafer et al. 1996), whereas hyperparameters for an XGBoost tree may involve learning rate, the minimum sum of weight of samples in a node, maximum tree depth, maximum number of leaves, minimum loss reduction, maximum delta step, the fraction of samples to be included in a tree, regularization term, and more[23]. Besides the number of possible hyperparameters, the range of hyperparameters is uncertain – some are bounded between 0 and 1, whereas others may have infinite possibilities. Trying each combination of the hyperparameters is computationally inefficient (and not necessary). The result of hyperparameter tuning presented in Appendix C (Table C2) is the best outcome of multiple trials of model cross-validation under a balanced accuracy (i.e., the arithmetic mean of sensitivity and specificity) maximization strategy with different configurations of hyperparameters and ranges. Default values of the hyperparameters that are not tuned are also presented in Table C2. Although the current configuration cannot guarantee that the model prediction results are the global optimal solutions, all model confusion matrices gave acceptable results (i.e., higher than the baseline before feature selection and tuning).

3.4.4 Model Performance

ML model performance reflects the extent to which the system of features as a whole is useful to classify the *paid* and *free* services. The results of ML models as measured by seven performance measures are presented in Table 3.10, which reflects model performance from different aspects (e.g., correctly assign the paid services with a *paid* label vs. the probability that a ML classifier will successfully classify a case in the right class). Except for the Naïve Bayes and Neural Network approaches, all other algorithms performed well and are comparable with each other (i.e., each approach shows the best performance in some of the measures). Since we have an imbalanced dataset (i.e., the number of *free* are larger than the number of *paid*), AUC, F-measure, balanced accuracy and Cohen's kappa are more unbiased and informative than other measures. Consequently, decision tree, gradient boost tree, and XG boost tree have the best overall performance. Although random forest, Adaboost and logistic regression achieve relatively lower scores, all measures are good enough to show the usefulness of the model. For the Neural Network algorithm, since we lack computational power to perform hyperparameter tuning, it is within expectation that a model with one hidden layer and up

[23] A full list of parameters can be found in XGBoost guide:
https://xgboost.readthedocs.io/en/latest/parameter.html

to 10 hidden neurons per layer does not perform well. These less-optimal results highlight the importance of model tuning for machine learning-based research.

Most models achieve an AUC above 0.95 (except for Naïve Bayes and Adaboost), suggesting excellent classification performance of the eleven features we have selected. It should be noted that logistic regression achieves an AUC of 1, which might be considered as the best performing method. However, logistic regression may be biased when the number of features is very small compared to the number of observations (Kumar et al. 2018). Overall, our predictive model with 11 selected features exhibits significant performance in classifying free-only services versus premium services, and the performance is consistent across different prediction methods.

Table 3.10 Classification Performance Comparison of Eight ML Algorithms

	LR	DT	GB	RF	ADA	XG	NB	NN
Recall	0.851	0.949	0.952	0.908	0.950	0.947	0.535	0.983
Precision	0.896	0.989	0.988	0.984	0.980	0.989	0.803	0.323
Specificity	0.956	0.995	0.995	0.993	0.991	0.995	0.941	0.070
F-measure	0.873	0.969	0.970	0.944	0.965	0.968	0.643	0.487
Balanced accuracy	0.903	0.972	0.973	0.951	0.971	0.971	0.738	0.527
AUC	1.000	0.988	1.000	0.988	0.942	0.992	0.882	0.662
Cohen's kappa	0.818	0.955	0.956	0.921	0.949	0.953	0.524	0.034

3.4.5 Interpreting Key Feature Rankings

Besides the overall performance of the feature bundles, feature importance is another useful tool to help physicians and the platform identify high-value consultation services. Not all ML algorithms provide importance indicators for each feature since the general purpose of ML is to estimate (learn) the mapping function from a bag of inputs to the output, so that we can apply the learned function to new cases in the future to make predictions. Herein, we only report the results from the four ML algorithms that provide some types of feature importance indicator.

Logistic regression and decision-tree algorithms are transparent in how a single feature influences the outcome. Gradient boost and random forest are decision-tree-based algorithms and provide feature importance indices[24] based on feature selection frequency as a decision node. Comparing the results from regression and tree-based analysis can provide complementary insights regarding how features combine to influence premium payment prediction. Table 3.11 presents the relative importance of features (ranked by the coefficient size for logistic regression, the level of split for the decision tree, and feature importance percentage for gradient boost tree and random forest). Whereas logistic regression shows the magnitude and direction of linear impacts, tree-based algorithms show the nested hierarchy of features, which is a type of non-linear relationship[25].

Table 3.11 Key Features Listed in Descending Order of Importance

	LR (coefficient)	DT (level of tree splits)	GB (importance%)	RF (importance %)
1	Response rate (-13.89***)	Offline connection (1)	Total Dialogue (30%)	Offline connection (24%)
2	Offline connection (-4.99***)	Social return (2)	Offline connection (30%)	PriorExam (20%)
3	Social return (-3.11***)	Total dialogue (2)	Response rate (25%)	Total dialogue (18%)
4	Patient posts (-2.63***)	Private (3)	Social return (8%)	Response rate (17%)
5	Total dialogue (2.47***)	Response rate (3)	Private (6%)	Patient post (9%)
6	PriorExam (1.70***)	PriorExam (4)	Patient posts	Social return (7%)
7	Private (-0.99***)	Answer_frq (4)	PriorExam	Private
8	Ranking 2 (-0.305***)	Patient posts (6)	Answer_frq	Answer_frq
9	Answer_frq (-0.14***)	Question_frq (6)	Question_frq	Question_frq
10	Question_frq (-0.13***)	Ranking 2 (8)	Title1	Title1
11	Title1 (-0.089***)	Title 1 (9)	Ranking2	Ranking2

Note. LR-logistic regression. Coefficients are shown in the brackets. DT – (simple) decision tree. The numbers in the brackets show the highest level of tree splits; GB- gradient boosted tree. The numbers in the brackets show the percentage feature importance (only those above 5% are shown); RF – random forest (only those above 5% are shown). The numbers in the brackets show the percentage feature importance.

In answer to research question two on the relative important of these features, in general, the four algorithms yield relatively consistent results in the top-ranked and lower-ranked features, whereas the ones in the middle are fuzzier. Offline connection, response rate,

[24] The explanation of feature importance calculation can be found in Friedman (2001), section 8.1.

[25] Note that a regularization procedure is applied to logistic regression. Thus, the large weight coefficients are penalized to avoid overfitting and to improve the generalization performance of the model.

social return, total dialogue, diagnoses from a prior exam and private status consistently rank high across four ML algorithms, whereas physician title, question frequency, and the second-tiered hospital ranking are consistently ranked low.

To address the third research question, we examine how these features interact in relation to patient payment. While using multiple ML approaches was useful for cross-validating the importance ranking of individual features, a tree structure is more appropriate for examining and interpreting the interactions between features because it explicitly displays the feature hierarchies and classification outcomes at each tree split. Thus, we address research question three using the tree structure provided in the Decision Tree algorithm. Figure 3.5 presents one randomly selected tree from one of the 10-fold cross-validation processes, and Table 3.12 describes the configuration of the top three level splits as well as our interpretation of the top feature configurations[26]. Top feature configurations are those that appear in the top three levels of the tree structure and for which one group's (free versus paid) cases account for a high enough percentage of the outcome to be useful for prediction. It should be noted that whereas coefficients for a regression suggest a symmetric and linear effect of independent variables (i.e., the features) on the dependent variable (e.g., the higher the total dialogue, the higher the probability of paying), the splits of a decision tree are hierarchically ranked which implies prior conditions for the impacts – for example, high total dialogue is the key feature that contributes to payment for offline returning patients, not online patients; for online patients, high total dialogue is useful for those who provided virtual gifts [*social return*].

Table 3.12 Decision Tree-Based Configuration of Feature Contributions

Top Feature Configurations	Dominant outcome	Our interpretation of the feature configurations
Offline	Free (84.2%)	Type 1- These configurations suggest a simple type of follow-up service resulting from previous offline diagnoses
Offline AND low total dialogue	Free (96.3%)	
Offline AND low total dialogue AND low response rate	Free (99%)	

[26] Note that the results for the other nine decision trees displayed similar structures including: the same four features always appear in the first four levels (even if the level on which they appear does change), the trees have similar depths (all between nine and eleven levels), and at the fourth level the number of records in each node are comparable.

Table 3.12 Decision Tree-Based Configuration of Feature Contributions		
Offline AND high total dialogue AND high response rate	Paid (100%)	Type 2- This configuration suggests a complex service extension from the previous offline diagnoses which requires intensive patient-provider interaction
Non-offline AND no social return	Paid (83.9%)	Type 3- These configurations suggest the patient has no offline connection with the physician but is paying premium for the online consultation rather than using an social return to show gratitude
Non-offline AND no social return AND non-private	Paid (97.6%)	
Non-offline AND social return AND low total dialogue	Free (91.3%)	Type 4- This configuration suggests the patient has no offline connection with the physician, has a less intensive online consultation experience, and offers a social returns as compensation instead of payment
Non-offline AND social return AND high total dialogue	Paid (71.6%)	Type 5- This configuration suggests the patient has no offline connection with the physician and engages in an intensive online interaction, providing both payment and a social return as compensation

Note. The sequence of features represents hierarchies in the tree. Only the top 4 levels are shown in the table and figure. Since free cases account for about 70% at baseline, we consider 80% as the threshold for a dominant *free* outcome and 70% as the threshold for a dominant *paid* outcome at the tree splits.

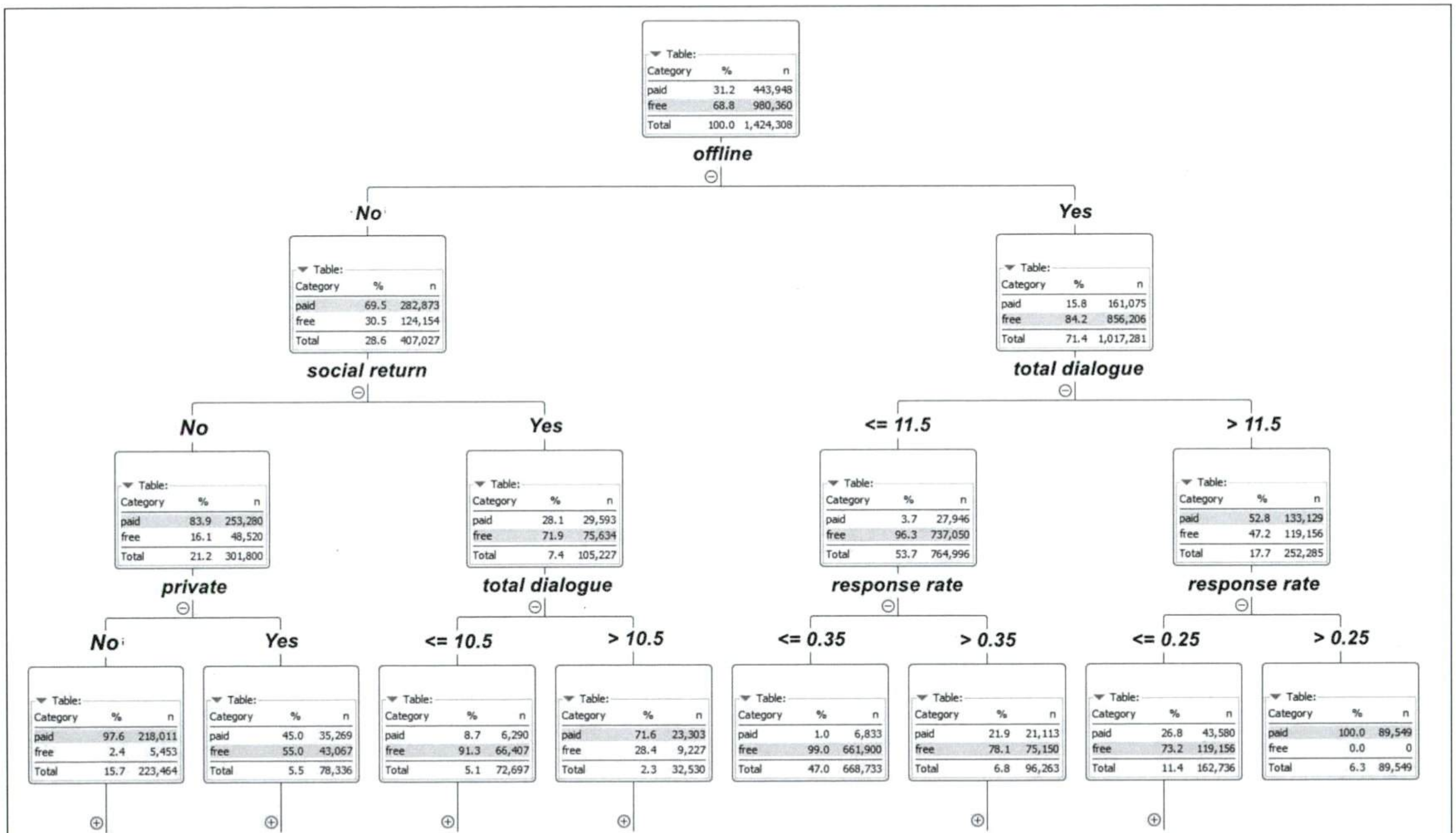


Figure 3.5 An Example of Decision Tree from One Round of 10-Fold Cross-Validation

From the top feature configurations listed above, we can observe that patients who have previous offline consultations with the physician are less likely to pay for premium service. This is different from our expectation. It is possible that these patients tend to take free opportunities to clarify the simple unsolved issues after their offline visits, as indicated by increasing the proportion of free services in the presence of low total dialogues and low response rate from the physicians (type 1). However, if complex issues emerge, they may still prefer to return to the offline healthcare channel rather than pay for premium service.

A second type of returning patients (type 2) may have complex issues and decide to stay online and pay. This represents a complex service extension which requires intensive patient-physician interaction. These returning patients with complex issues may communicate more and expect fast physician responses (type 2). Thus, those returning patients who frequently communicate with the physicians (probably due to the complexity of the issue) and receive intensive responses are associated with high probability of payment. This seems to indicate that one potential differentiator between returning patients who decide to pay on the platform is the level of detail in the interaction and the speed of physician response.

For online patients who have no prior connection with the physician, those who do not provide social returns (e.g., thank-you letters and virtual gifts) seem more likely to pay (69.5% paid consultations, type 3). There may be a psychological compensation effect (Bäckman and Dixon 1992) where giving virtual gifts substitutes for the actual payment and balances the sense of "guilt" after receiving free services. However, in cases where the service between patients and physicians with no offline connection is intense (i.e., large amount of dialogue), patients provide both virtual gifts and premium payment to show their appreciation (type 4 verses type 5).

The high-level presence of *private* in one of the tree branches deserves more attention. One of the major reasons that patients do not use online healthcare service is privacy concerns (Angst and Agarwal 2009; Bansal and Gefen 2010). *Privacy* represents a function provided by the platform which allows patients to set their dialogues as private so they cannot be viewed by other people. Patients who enable this function may have a higher privacy concern than those who do not. While the physician can still see health information about the patient, they cannot see the patient's previous consultations with other physicians on the platform. This may limit the information available to the physician and decrease the usefulness of their diagnosis and recommendation. Since online medical consultation inevitably requires patients to reveal sensitive health-related information, patients who have high privacy concerns may not be deeply involved in the online consultation and may receive less useful diagnosis, thus are less likely to pay after the initial free interactions.

It is important to note that there two features that do not appear in the top three levels of the tree structure, but which are consistently highly ranked by the ML algorithms. Based on logistic regression, *PriorExam* has positive impacts on payments, and two of the other ML algorithms rank this feature highly as well. It is possible that some patients use the platform to look for a second opinion from a different physician. If patients are seeking a second opinion (rather than looking for initial treatments or diagnoses), the convergence or divergence of the two opinions may influence the payment decision. It may be an opportunity for the platform to provide supportive IT functionalities that intervene in the opinion comparison process. In addition, results from logistic regression and random forest also rank the *number of patient posts* as an important predictor. This is in line with the highly ranked total dialogues (i.e., patients posts are part of total dialogue), which implies that the patient involvement during the consultation is an important factor in getting paid.

Surprisingly, physicians' offline reputation, such as title and the affiliated hospital ranking, does not rank highly in the ML algorithms and does not appear in the top three levels of the tree structure. These physician offline reputation features are frequently employed by previous studies in similar contexts. Although our logistic regression results exhibit significant coefficients for these offline reputation features, in the tree structure they only play a role in combination with other features in the lower levels. Thus, our results show that the regression approach may not be the best method to detect the impacts of various predictors and may yield over-simplified interpretation – regression only shows a linear additive relationship and excludes collinearity, whereas in reality, complex interactions and equifinality may play significant roles.

In summary, the source of patients (offline returning or online directly) seems to be a key differentiator for payment, which may be due to the different motivations and service requirements inherent in these two types of patients. In line with previous studies, physicians' offline reputation (e.g., title and affiliation) is a key predictor. However, our results show that it is not as important as patient-physician interaction (e.g., total dialogues, response rate, patient posts). Privacy setting and social return, two features pertaining to the platform functionality, play important roles as well.

3.4.6 Machine Learning Model Verification

In the previous subsections, we have tested the model using multiple classifiers and the 10-fold cross-validation technique to evaluate its classification performance. In order to ensure that the data was representative of the population on the online medical consultation platform, we attempted to minimize human input into the selection process. However, there are some particularities of our data that could lead these results to be biased toward certain outcomes. Thus, to guarantee the robustness of the model, we perform several additional tests to verify the impacts of potential variation sources. The purpose of model verification, as opposed to model testing, is to see if the ML model will behave consistently under different circumstances. We focus on four aspects:

(1) **Perturbance due to imbalanced data**: In our dataset, there are significantly more free trial patients than paying patients. When training with an imbalanced dataset, the standard ML classifiers are often more sensitive to detecting the majority class patterns, thus increasing the possibility that the minority cases are misclassified and raising a type II error (López et al. 2013). In practice, incorrectly categorizing a potential premium subscriber to the free-trial-only consumer group may cost more for the business to establish a long-term consumer relationship. In our major analysis, we used widely discussed remedy techniques such as boosting, bagging and repeated random resampling, and less biased measures to evaluate the performance (He and Garcia 2008; Khalilia et al. 2011). Although the imbalance level in our dataset is not high (approximately 1:2.2), to

examine whether our classification based on imbalanced data suffered from a type II error, we re-run the analysis using an under-sampling technique (i.e., reduce the number of free trial patients to balance the groups) to verify the impacts of such perturbance.

(2) **Geographic differences** (i.e., areas with rich healthcare resources versus areas with few resources): Allowing patients from remote areas to access to high-quality healthcare resources without traveling far is one of the advantages of digital healthcare service delivery. Although online patients can select their healthcare service provider without considering the geographic constraints, many patients would still choose a physician near them, so that if offline diagnoses or treatments are needed, they can easily transfer to an accessible hospital. Considering the possible systematic differences between healthcare resource-rich areas and the remote areas with few resources, we compare and verify our model by splitting the dataset based on service providers' location.

(3) **The rapid growth period:** From Figure 3.1 (the trend diagram) we can observe that in addition to 2009, there is a second rapid growth period after 2015. National policies and the increased acceptance of online medical consultation by the general population may facilitate such growth. Considering the potential systematic differences in consumer behaviors before and after, we perform additional analysis with balanced four-year data (2015 to 2018) to examine whether the model performance is significantly influenced by these unknown market- or national-level disturbances.

(4) **Outliers**: Fraudulent cases may impact the performance of an ML system, and many ML algorithms are sensitive to outliers. For example, logistic regression can be easily impacted by extreme cases in the training data, and AdaBoost may increase the weights of misclassified outliers. In the main analysis, we exclude many outliers using a 95% quantile as a threshold, and the results show that the model performs well in a "normal" situation. To examine whether our model is robust under those extreme or even fraudulent cases (due to web scraping errors), we perform additional analysis with a balanced dataset, including randomly selected outliers.

The detailed results of these ML model verification are presented in Appendix D. The results of four analyses indicate that our model is robust to sample distribution (e.g., imbalances, classes and outliers) and potential systematic differences (e.g., geographic location and market changes), as indicated by minor changes in the model performance measures. Moreover, although the specific feature ranking may change, the overall trends remain the same – offline connection, service intensity and patient involvement features are ranked highly, whereas physician offline reputation is ranked relatively low.

3.5 Discussion

The freemium business model is a potentially useful strategy for attracting patients who are reluctant to pay for digital healthcare services. In this study, we focus on online medical consultation, a type of emerging digital healthcare service that has received much attention in recent years. Our objective is to understand the observable features of online medical consultation services that may contribute to premium payment so that the business can identify high-value services and take actions to better manage service providers and their offerings. We started with a feature selection procedure to ensure the relative parsimony of our model. Eight algorithms that represent different ML philosophies are used to cross-validate the predictive power of the model. As an initial study using machine learning approaches to identify key service-related features and to make predictions, we do not aim to incrementally improve prediction accuracy by engineering the features or developing new algorithms. Rather, our goal is to minimize overfitting and develop a generalizable predictive model that both has sufficient explanatory power and succinct enough that it can be used by general data consumers, such as medical consultation platforms and service providers. Thus, we use classic ML algorithms and try to be transparent in model tuning and cross-validation procedures.

3.5.1 Principle Findings

After the data-driven feature selection with 18 initial service features, we identified 11 important ones that can cover multiple aspects of online consultation while keeping the analytical model relatively parsimonious. The high performance across the ML algorithms (except for the Naïve Bayes and the Neural Network) demonstrates that our 11-feature model is a useful predictive tool. The platform or service providers can leverage these service features to better manage healthcare resources and service delivery efforts (RQ1).

In terms of feature importance (RQ2 and RQ3), our results show that although physician reputation is important, service quality and patient involvement appear to contribute more to premium payment. We further identified five subgroups of patient-provider relationships based on the decision tree structure. The feature configurations show how patient characteristics, service involvement by both physicians and patients, and physician characteristics interact to result in different payment outcomes. These configurations highlight offline connection as a key differentiator for premium payment versus free-trial only services. In addition, responsive interactions that encourage multiple timely responses by patients seem to facilitate payment.

3.5.2 Comparison with Prior Work

As has been discussed, previous studies have identified various antecedents of payment in the context of online medical consultation. The majority of service features included in our predictive model were examined in existing research, while several new features specific to this type of platform were also investigated. As opposed to previous work which highlights the important role of physician reputation such as indicated by title and affiliation (e.g., Deng et al. 2019; Guo et al. 2017; Li et al. 2019a), the feature ranking in our model emphasized the more important role of patient source and patient-provider interactions during the service as key drivers of payment. According to existing freemium research, it is likely that patients experience different stages of awareness and learning during the phases of physician selection, free service and paid service (Liu et al. 2014). Whereas physician reputation may increase patients' initial service awareness and influence physician selection, it seems that direct positive service experience (i.e., service quality and intensive involvement) ensures further commitment.

In contrast to previous results which show the positive influence of social ties between physician and patients on payment (Guo et al. 2018), our results show that a prior offline relationship with the physician (which implies a stronger social tie than no offline relationship) does not always seem to be a facilitating factor for premium payment online (i.e., patient type I from Table 3.12). The design of the platform (e.g., the "check-in" function for returning patients) and freemium rules (e.g., three initial free online services) seem to facilitate the development of stronger ties. However, these patients may only take advantage of these features and rules to acquire free services for simple follow-ups without devoting themselves in establishing stronger online relationships with the physicians. Thus, the patients may prefer to visit their physicians offline for solving more complex issues, avoiding payment online. This may also be due to the patients' motivation for consulting healthcare issues online — patients may be willing to try online consultation to solve non-complex issues since it is a time and cost effective approach, but it may be difficult for patients to completely shift their healthcare practices and habits since the associated risk can be high, and sufficient knowledge regarding this new type of service is needed for patients to trust and accept it (Chernof et al. 1999). Accordingly, we observe that when patients are facing complex healthcare issues which require intensive communication, and when the physician's response rate is high enough, these returning patients are likely to pay (i.e., patient type II). Otherwise, they will only take a chance on free services without further engagement.

Previous studies also highlighted virtual gifts as a positive signal for payment (Yang et al. 2015; Yang and Zhang 2019). However, our findings suggest that virtual gift on the platform may be a double-edged sword. For patients who have no prior offline connections with the physician, allowing them to show gratitude with a virtual gift function may not be a good strategy as this type of patient may substitute this virtual gift for payment (i.e., patient type IV). However, if the service is intensive, virtual gifts and payment will be additive rather than virtual gifts replacing payment (i.e., patient types III and V).

In summary, in line with previous literature on online service delivery, responsive service is a key antecedent of payment (Berry et al. 1994; Storey et al. 2016; Wakefield and Blodgett 1999). However, in our context, responsive service seems to be more likely to attract premium payment for online patients, but not offline returning patients. Encouraging patient engagement (e.g., encouraging multiple timely interactions with the physician) may help promote payment as well. Since each response to the physician counts as one free-trial for the patients, reluctance to consult further may arise at the end of each conversation turn. Persuading patients to keep on responding in a timely manner should be beneficial for establishing long-term patient-physician collaboration and attracting payments. The feature ranking and configuration results from four ML approaches (i.e., LR coefficient ranking, feature hierarchies in DT, and feature importance score ranking in RF and GB) indicate that these observable service features are not generating linear impacts, a finding which was not evident in previous studies using linear regression approaches. Although tree-based approaches provide useful implications on how features are ranked hierarchically, further investigations are needed to reveal the magnitude and provide finer-grained explanations.3.5.3 Contributions

This study contributes to research and practice in several ways. First, the high-ranked features highlight the importance of consumer-provider relationships as a key factor in facilitating premium payment, which is seldom examined in previous freemium or online medical consultation research. Whereas previous freemium research generally provides an individual-level explanation for consumers' willingness to pay (e.g., psychological factors or cognitive evaluation of the service quality), our results show that the channel from which the consumers come (e.g., from offline, or finding a physician online directly) may largely determine their online behaviors in this context.

Second, our study extends the existing research on online medical consultation by showing the role of physician-patient interaction. Other than Guo and colleagues (2018) who examine the influence of strong versus weak physician-patient ties (as indicated by the amount of interaction), most existing medical online consultation research does not investigate the importance of the physician-patient interaction. Our work goes further and demonstrates that service providers' interactions with the consumers (e.g., response rate) and consumer involvement (e.g., communication intensity) also largely influence their payment. This matches the experiential nature of online consultation, where not only observational learning (e.g., examining other peoples' reviews) but also the direct experience has significant impacts on payment decisions.

Third, our study also makes practical contributions to the management of online medical consultation and digital healthcare delivery in general – platform owners and administrators can utilize our model to identify and promote high-value services and service providers, which in turn should help support the long-term success of the platform. For example, the platform can develop different governance mechanisms that draw on our system level view (for example, separate collection of mechanisms for returning patients and online-only patients). Whereas enhancing patient engagement and physician response rate are particularly important for offline returning patients – perhaps due to their offline service experience as a comparison – ensuring privacy and avoiding alternative channels to cannibalize monetary return are the keys for switching online-only patients to premium subscribers.

Fourth, our study may have broader implications for the distribution of healthcare resources online. In an offline setting, distribution of healthcare resources usually falls under the responsibility of the region's or country's healthcare system. In an online environment, "celebrity" effects and a lack of regulation may lead to inefficient digital healthcare resource distribution. For example, in our data physicians with a critical mass of visible patient-physician consultations may become popular and attract a large number of online patients, while other equally qualified physicians without a critical mass of existing interactions may have trouble attracting patients. The capacity of popular physicians may become oversaturated by too many online appointments, resulting in slow responses and unsatisfactory service delivery. This could impact the willingness of patients to continue using online medical consultation. There may also be a dampening effect on equally qualified physicians with fewer previous interactions where inability to attract patients hinders their enthusiasm to participate in online medical consultation. As a result, inefficient healthcare resource distribution online will create a new type of inequity for both patients and physicians.

Our study also makes methodological contributions. Machine learning has been increasingly adopted by IS researchers as an emerging methodology to learn from data, build predictive models and show the importance of explanatory features when the research question is associated with complex explanation and a large amount of data

(Adamopoulos et al. 2018; Gong et al. 2018; Li and Qin 2017; Lukyanenko et al. 2019; Wang et al. 2018). We aim to provide a rigorous analysis procedure along with a transparent report so that future research can easily follow or adapt our methodological steps. For example, previous IS research seldom explicates their hyperparameter tuning (or does not have such a procedure) despite the importance of this prior step in determining prediction performance. The measures and configurations used in this study can be a reference list for future researchers who want to conduct comprehensive model tuning. Also, our model training and validation procedures cover most of the applicable ML algorithms for a binary classification question, except for several algorithms that do not fit our research questions (e.g., Support Vector Machine algorithm and k-Nearest Neighbors algorithm). Future research that is conducted in similar contexts can compare their results with ours as an external validation procedure to further examine model efficiency and generalizability.

Moreover, previous studies in similar contexts are dominated by regression-based analysis (e.g., Cao et al. 2017; Li et al. 2019a; Yang et al. 2015; Yang and Zhang 2019). Whereas such traditional statistical modeling is useful for formalizing the relationship between the variables, it requires that a number of assumptions be met (e.g., linear assumption, independence of observations, normally distributed errors, the type of data, homoscedasticity). Consequently, it might be difficult to use traditional statistical modeling approaches to make predictions or formalize relationships from massive amounts of fine-grained consumer behavior data. This study replicates statistical modeling procedures to a great extent (e.g., creating variables that capture the key explanatory factors, estimating the model, interpreting the relationship between explanatory variables and the outcome). However, instead of focusing on coefficient estimation, which is the key objective of statistical modeling, the ML approach helps us understand the relationships between the system of inputs (e.g., fine-grained consumer data, or features) and outcome from a more holistic perspective.[27] For example, in our

[27] A more holistic perspective is needed to capture the complex interactions. To illustrate, a negative coefficient from the logistic regression does not necessarily mean a truly negative impact across the whole explanatory sphere if other factors are taken into consideration (e.g., having an offline connection does not always have negative impacts).

study, logistic regression and tree-based ML classifiers provide complementary views on complex interactions among the relevant service features that contribute to payment, in which various factors are nested and interact with each other in non-linear ways. Our methods, as well as the complementary results from logistic regression and tree-based analysis, show the usefulness of mining finer-grained consumer data and the importance of considering non-linear relationships and interactions between factors in future research.

3.5.4 Limitations

We have also identified several limitations. First, since the model analysis was based on the results of feature selection, the subjective decision making during the feature selection procedure may cause some bias in the subsequent analysis. Although the current results achieved satisfactory overall performance, a different set of features that are comparable with the current ones can be tried to cross-validate our model. For example, during high-correlation filtering, the *other* feature highly correlated with our currently selected feature could be selected instead and the analysis could be rerun. Alternatively, more restrictive criteria can be used when combining the results from the four feature selection approaches.

Second, due to the lack of computational resources, we performed our hyperparameter tuning within a limited range, which cannot ensure the identification of globally optimal solutions. This is especially true for neural network analysis that needs a high-performance computation to train the model and find the best solution. However, our current solution is still satisfactory since all performance measures reach good levels (except for Neural Networks).

The third limitation of our study is that only premium subscriptions with short-term services are included (e.g., 48-hour text-based consultation, one-time phone call), whereas long-term contracts such as private doctor, family doctor and expert team services cannot be easily identified with the currently available data[28]. However, patients' decision-making criteria for adopting a long-term contract can be very different from those for a

[28] This data is not part of the current dataset and would take several months to collect and clean. This work is currently underway for future research on this online medical consultation platform.

one-time or repetitive payment. Future research can conduct further empirical examinations of the service characteristics that may influence long-term premium subscription.

3.5.5 Avenues for Future Research

Our work also provides several opportunities for future research. First, Future online medical consultation research should investigate how the development of patient-provider relationships influences patients' premium payment over time (e.g., by analyzing the communication content). For research on freemium issues in the related fields such as the online knowledge market or experience goods market in general, investigating the nature and trajectory of consumer-service provider relationship development can be an interesting angle to extend existing understanding of how freemium businesses succeed.

Second, although our results suggest that offline connection does not seem to contribute to online payment, future research should examine if it brings additional offline benefits for the physicians. For example, offering free online follow-up may reduce congestion in physicians' offline waiting lists and reduce the time and resources associated with a face-to-face follow-up appointment (e.g., physician time in-office; administrative overhead to schedule, confirm, and process appointments; congested waiting rooms). When offline patients switch to online delivery for follow-up, the timing of the interaction is flexible, and physicians can attend to the online follow-up at a time that does not interrupt the flow of their other activities. This may help physicians free up their time and resources for new cases and patients with more complex issues. Future research may investigate the impacts of offering free online follow-up medical consultation on a physician's offline benefits.

Future research should also examine the inequity in healthcare resource distribution that may develop for both patients and physicians on online medical consultation platforms as well as how the platforms can used to alleviate this inequity. Platform technologies can be used as digital empowerment (Leong et al. 2016; Mäkinen 2006; Zhao et al. 2008) to enrich participation by physicians and patients. For example, the platform can serve as a conversation moderator. Our results indicate the importance of patients' timely and repetitive interaction with the physicians, however the rules of freemium may create constraints on these interactions (e.g., limit of three free trials in our data). The platforms could be designed to increase and expand the flow of information (e.g., probing, automatically requesting new information, conciliating waiting patients) to increase patients' feeling that they are engaged in an ongoing interactive conversation. Further research should examine these and other technology-mediated approaches to unfold the process and impacts of such digital empowerment in digital healthcare delivery.

3.6 Conclusion

Healthcare delivery is stepping into a new digital era. Online delivery approaches such as online medical consultation are increasingly common, and these platforms open a new front in which individual consumers and independent service providers play increasingly important roles. These multisided platforms can also utilize different business models such as freemium. While business model innovation in knowledge-intensive industries has received much research attention in the context of inter- or intra-organizational collaboration (Hertog 2000; Muller and Doloreux 2009; Rajala and Westerlund 2007), few studies have examined the freemium model in a multisided platform economy. Online medical consultation is one such multisided platform that connects patients and physicians through complex interactions.

Online medical consultation gives patients a new channel and more options to access healthcare services. However, reluctance and concerns exist which may impede patients' physician selection and payment on the platform. The current study explores the key service features that may contribute to patient payment under a freemium business model. By mining massive consultation data using machine learning approaches, our results highlight the important roles of patient source (i.e., offline returning patients vs. online patients), service delivery quality (i.e., service intensity, physician responsiveness), patient involvement (i.e., providing social returns, revealing prior examinations, number of patient posts), and physician reputation in facilitating premium payment.

References

Abbasi, A., and Chen, H. 2008. "Writeprints: A Stylometric Approach to Identity-Level Identification and Similarity Detection in Cyberspace," *ACM Transactions on Information Systems (TOIS)* (26:2), p. 7.

Abu-Salim, T., Onyia, O. P., Harrison, T., and Lindsay, V. 2017. "Effects of Perceived Cost, Service Quality, and Customer Satisfaction on Health Insurance Service Continuance," *Journal of Financial Services Marketing* (22:4), pp. 173-186.

Adamopoulos, P., Ghose, A., and Todri, V. 2018. "The Impact of User Personality Traits on Word of Mouth: Text-Mining Social Media Platforms," *Information Systems Research* (29:3), pp. 612-640.

Angst, C. M., and Agarwal, R. 2009. "Adoption of Electronic Health Records in the Presence of Privacy Concerns: The Elaboration Likelihood Model and Individual Persuasion," *MIS quarterly* (33:2), pp. 339-370.

Anselmsson, J., Johansson, U., and Persson, N. 2007. "Understanding Price Premium for Grocery Products: A Conceptual Model of Customer-Based Brand Equity," *Journal of Product & Brand Management* (16:6), pp. 401-414.

Bäckman, L., and Dixon, R. A. 1992. "Psychological Compensation: A Theoretical Framework," *Psychological Bulletin* (112:2), pp. 259-283.

Bansal, G., and Gefen, D. 2010. "The Impact of Personal Dispositions on Information Sensitivity, Privacy Concern and Trust in Disclosing Health Information Online," *Decision support systems* (49:2), pp. 138-150.

Batista, G. E., Prati, R. C., and Monard, M. C. 2004. "A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data," *ACM SIGKDD explorations newsletter* (6:1), pp. 20-29.

Bawa, K., and Shoemaker, R. 2004. "The Effects of Free Sample Promotions on Incremental Brand Sales," *Marketing Science* (23:3), pp. 345-363.

Berry, L. L., Parasuraman, A., & Zeithaml, V. A. 1994. "Improving service quality in America: lessons learned," *Academy of Management Perspectives*, (8:2), pp. 32-45.

Bhargava, H. K., and Choudhary, V. 2008. "Research Note—When Is Versioning Optimal for Information Goods?," *Management Science* (54:5), pp. 1029-1035.

Biesdorf, S., and Niedermann, F. 2014. "Healthcare's Digital Future." *McKinsey Insights* Retrieved May 05, 2019, from https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/healthcares-digital-future

Brodersen, K. H., Ong, C. S., Stephan, K. E., and Buhmann, J. M. 2010. "The Balanced Accuracy and Its Posterior Distribution," *2010 20th International Conference on Pattern Recognition*: IEEE, pp. 3121-3124.

Callahan, E. J., Hilty, D. M., and Nesbitt, T. S. 1998. "Patient Satisfaction with Telemedicine Consultation in Primary Care: Comparison of Ratings of Medical and Mental Health Applications," *Telemedicine Journal* (4:4), pp. 363-369.

Cao, X., Liu, Y., Zhu, Z., Hu, J., and Chen, X. 2017. "Online Selection of a Physician by Patients: Empirical Study from Elaboration Likelihood Perspective," *Computers in Human Behavior* (73), pp. 403-412.

CBC. 2017. "Controversial for-Fee Service Connects Ontario Doctors and Patients Via Video Chat." *CBC News* Retrieved May 15th, 2019, from https://www.cbc.ca/news/canada/london/maple-health-online-doctor-visits-1.4194285

Chada, B. V. 2017. "Virtual Consultations in General Practice: Embracing Innovation, Carefully," *The British Journal of General Practice* (67:659), p. 264.

Chandukala, S. R., Dotson, J. P., and Liu, Q. 2017. "An Assessment of When, Where and under What Conditions in-Store Sampling Is Most Effective," *Journal of Retailing* (93:4), pp. 493-506.

Chesbrough, H. 2010. "Business Model Innovation: Opportunities and Barriers," *Long range planning* (43:2-3), pp. 354-363.

Chernof, B. A., Sherman, S. E., Lanto, A. B., Lee, M. L., Yano, E. M., & Rubenstein, L. V. 1999. “Health habit counseling amidst competing demands: effects of patient health habits and visit characteristics”. *Medical care*, 37(8), 738-747.

Cooil, B., Winer, R. S., and Rados, D. L. 1987. "Cross-Validation for Prediction," *Journal of Marketing Research* (24:3), pp. 271-279.

Cordina, J., Jones, E. P., Kumar, R., and Martin, C. P. 2018. "Healthcare Consumerism 2018: An Update on the Journey." Retrieved May 05, 2019, from https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/healthcare-consumerism-2018

Cox, J. 2017. "Play It Again, Sam? Versioning in the Market for Second - Hand Video Game Software," *Managerial and Decision Economics* (38:4), pp. 526-533.

Cui, G., Leung Wong, M., Zhang, G., and Li, L. 2008. "Model Selection for Direct Marketing: Performance Criteria and Validation Methods," *Marketing Intelligence & Planning* (26:3), pp. 275-292.

Dash, M., and Liu, H. 1997. "Feature Selection for Classification," *Intelligent data analysis* (1:1-4), pp. 131-156.

Datta, H., Foubert, B., and Van Heerde, H. J. 2015. "The Challenge of Retaining Customers Acquired with Free Trials," *Journal of Marketing Research* (52:2), pp. 217-234.

Deloitte. 2019. "2019 Us and Global Health Care Industry Outlook." Retrieved May 05, 2019, from https://www2.deloitte.com/us/en/pages/life-sciences-and-health-care/articles/us-and-global-health-care-industry-trends-outlook.html

Deng, Z., Hong, Z., Zhang, W., Evans, R., and Chen, Y. 2019. "The Effect of Online Effort and Reputation of Physicians on Patients' Choice: 3-Wave Data Analysis of China's Good Doctor Website," *Journal of medical Internet research* (21:3), p. e10170.

Dey, D., and Lahiri, A. 2016. "Versioning: Go Vertical in a Horizontal Market?," *Journal of Management Information Systems* (33:2), pp. 546-572.

Dey, D., Lahiri, A., and Liu, D. 2013. "Consumer Learning and Time-Locked Trials of Software Products," *Journal of Management Information Systems* (30:2), pp. 239-268.

Dick, P. T., Filler, R., and Pavan, A. 1999. "Participant Satisfaction and Comfort with Multidisciplinary Pediatric Telemedicine Consultations," *Journal of pediatric surgery* (34:1), pp. 137-142.

Dulleck, U., and Kerschbamer, R. 2006. "On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods," *Journal of Economic literature* (44:1), pp. 5-42.

El-Manstrly, D. 2016. "Enhancing Customer Loyalty: Critical Switching Cost Factors," *Journal of Service Management* (27:2), pp. 144-169.

Foubert, B., and Gijsbrechts, E. 2016. "Try It, You'll Like It—or Will You? The Perils of Early Free-Trial Promotions for High-Tech Service Adoption," *Marketing Science* (35:5), pp. 810-826.

Friedman, J. H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of statistics*), pp. 1189-1232.

Gallaugher, J. M., and Wang, Y.-M. 1999. "Network Effects and the Impact of Free Goods: An Analysis of the Web Server Market," *International Journal of Electronic Commerce* (3:4), pp. 67-88.

García, S., Fernández, A., Luengo, J., and Herrera, F. 2009. "A Study of Statistical Techniques and Performance Measures for Genetics-Based Machine Learning: Accuracy and Interpretability," *Soft Computing* (13:10), p. 959.

GarnerInsights. 2018. "Online Doctor Consultation Market: Global Market Synopsis, Growth Factors, Industry Segmentation, Regional Analysis and Competitive Analysis 2017 - 2025." Retrieved May 05, 2019, from http://garnerinsights.com/Online-Doctor-Consultation-Market-Global-Market-Synopsis-Growth-Factors-Industry-Segmentation-Regional-Analysis-And-Competitive-Analysis-2017---2025

Gong, J., Abhishek, V., and Li, B. 2018. "Examining the Impact of Keyword Ambiguity on Search Advertising Performance: A Topic Model Approach," *MIS Quarterly* (42:3), pp. 805-829.

Greenhalgh, T., Vijayaraghavan, S., Wherton, J., Shaw, S., Byrne, E., Campbell-Richards, D., Bhattacharya, S., Hanson, P., Ramoutar, S., and Gutteridge, C. 2016. "Virtual

Online Consultations: Advantages and Limitations (Vocal) Study," *BMJ open* (6:1), p. e009388.

Guo, S., Guo, X., Fang, Y., and Vogel, D. 2017. "How Doctors Gain Social and Economic Returns in Online Health-Care Communities: A Professional Capital Perspective," *Journal of Management Information Systems* (34:2), pp. 487-519.

Guo, S., Guo, X., Zhang, X., and Vogel, D. 2018. "Doctor–Patient Relationship Strength's Impact in an Online Healthcare Community," *Information Technology for Development* (24:2), pp. 279-300.

Hall, M. A. 2000. "Correlation-Based Feature Selection of Discrete and Numeric Class Machine Learning," University of Waikato, Department of Computer Science.

Hamari, J., Hanner, N., and Koivisto, J. 2017. "Service Quality Explains Why People Use Freemium Services but Not If They Go Premium: An Empirical Study in Free-to-Play Games," *International Journal of Information Management* (37:1), pp. 1449-1459.

Hansen, S. W., Gogan, J. L., Baxter, R. J., and Garfield, M. J. 2019. "Informed Collaboration in Health Care: An Embedded - Cases Study in Geriatric Telepsychiatry," *Information Systems Journal* (29:2), pp. 514-547.

He, H., and Garcia, E. A. 2008. "Learning from Imbalanced Data," *IEEE Transactions on Knowledge & Data Engineering*:9), pp. 1263-1284.

Heiman, A., McWilliams, B., Shen, Z., and Zilberman, D. 2001. "Learning and Forgetting: Modeling Optimal Product Sampling over Time," *Management Science* (47:4), pp. 532-546.

Hertog, P. d. 2000. "Knowledge-Intensive Business Services as Co-Producers of Innovation," *International journal of innovation management* (4:04), pp. 491-528.

Huang, H.-C. 2016. "Freemium Business Model: Construct Development and Measurement Validation," *Internet Research* (26:3), pp. 604-625.

Jain, D., Mahajan, V., and Muller, E. 1995. "An Approach for Determining Optimal Product Sampling for the Diffusion of a New Product," *Journal of Product Innovation Management* (12:2), pp. 124-135.

Jeni, L. A., Cohn, J. F., and De La Torre, F. 2013. "Facing Imbalanced Data--Recommendations for the Use of Performance Metrics," *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*: IEEE, pp. 245-251.

Jiang, S. 2019. "The Relationship between Face-to-Face and Online Patient-Provider Communication: Examining the Moderating Roles of Patient Trust and Patient Satisfaction," *Health Communication* (online preprint), pp. 1-9.

Jiang, Z., and Sarkar, S. 2009. "Speed Matters: The Role of Free Software Offer in Software Diffusion," *Journal of Management Information Systems* (26:3), pp. 207-240.

Jiao, Y., and Du, P. 2016. "Performance Measures in Evaluating Machine Learning Based Bioinformatics Predictors for Classifications," *Quantitative Biology* (4:4), pp. 320-330.

Jin, J., Yan, X., Li, Y., and Li, Y. 2016. "How Users Adopt Healthcare Information: An Empirical Study of an Online Q&a Community," *International journal of medical informatics* (86), pp. 91-103.

Kannan, P., and Kopalle, P. K. 2001. "Dynamic Pricing on the Internet: Importance and Implications for Consumer Behavior," *International Journal of Electronic Commerce* (5:3), pp. 63-83.

Khalilia, M., Chakraborty, S., and Popescu, M. 2011. "Predicting Disease Risks from Highly Imbalanced Data Using Random Forest," *BMC medical informatics and decision making* (11:1), p. 51.

Kitchens, B., Dobolyi, D., Li, J., and Abbasi, A. 2018. "Advanced Customer Analytics: Strategic Value through Integration of Relationship-Oriented Big Data," *Journal Of Management Information Systems* (35:2), pp. 540-574.

Kumar, N., Venugopal, D., Qiu, L., and Kumar, S. 2018. "Detecting Review Manipulation on Online Platforms with Hierarchical Supervised Learning," *Journal of Management Information Systems* (35:1), pp. 350-380.

Kumar, V. 2014. "Making" Freemium" Work," *Harvard business review* (92:5), pp. 27-29.

Lammers, H. B. 1991. "The Effect of Free Samples on Immediate Consumer Purchase," *Journal of Consumer Marketing* (8:2), pp. 31-37.

Larsen, K. R., and Bong, C. H. 2016. "A Tool for Addressing Construct Identity in Literature Reviews and Meta-Analyses," *Mis Quarterly* (40:3), pp. 529-551.

Lee-Wingate, S. N., and Corfman, K. P. 2010. "A Little Something for Me and Maybe for You, Too: Promotions That Relieve Guilt," *Marketing Letters* (21:4), pp. 385-395.

Leong, C. M. L., Pan, S.-L., Newell, S., and Cui, L. 2016. "The Emergence of Self-Organizing E-Commerce Ecosystems in Remote Villages of China: A Tale of Digital Empowerment for Rural Development," *MIS Quarterly* (40:2), pp. 475-484.

Li, J., Tang, J., Jiang, L., Yen, D. C., and Liu, X. 2019a. "Economic Success of Physicians in the Online Consultation Market: A Signaling Theory Perspective," *International Journal of Electronic Commerce* (23:2), pp. 244-271.

Li, J., Zhang, Y., Ma, L., and Liu, X. 2016. "The Impact of the Internet on Health Consultation Market Concentration: An Econometric Analysis of Secondary Data," *Journal of medical Internet research* (18:10), p. e276.

Li, X.-B., and Qin, J. 2017. "Anonymizing and Sharing Medical Text Records," *Information Systems Research* (28:2), pp. 332-352.

Li, Y., Ma, X., Song, J., Yang, Y., and Ju, X. 2019b. "Exploring the Effects of Online Rating and the Activeness of Physicians on the Number of Patients in an Online Health Community," *Telemedicine and e-Health*).

Liu, C. Z., Au, Y. A., and Choi, H. S. 2014. "Effects of Freemium Strategy in the Mobile App Market: An Empirical Study of Google Play," *Journal of Management Information Systems* (31:3), pp. 326-354.

Liu, J., Bian, Y., Ye, Q., and Jing, D. 2018. "Free for Caring? The Effect of Offering Free Online Medical-Consulting Services on Physician Performance in E-Health Care," *Telemedicine and e-Health*).

Liu, X., Guo, X., Wu, H., and Wu, T. 2016. "The Impact of Individual and Organizational Reputation on Physicians' Appointments Online," *International Journal of Electronic Commerce* (20:4), pp. 551-577.

López, V., Fernández, A., García, S., Palade, V., and Herrera, F. 2013. "An Insight into Classification with Imbalanced Data: Empirical Results and Current Trends on Using Data Intrinsic Characteristics," *Information sciences* (250), pp. 113-141.

Lu, N., and Wu, H. 2016. "Exploring the Impact of Word-of-Mouth About Physicians' Service Quality on Patient Choice Based on Online Health Communities," *BMC medical informatics and decision making* (16:1), p. 151.

Lukyanenko, R., Parsons, J., Wiersma, Y. F., and Maddah, M. 2019. "Expecting the Unexpected: Effects of Data Collection Design Choices on the Quality of Crowdsourced User-Generated Content " *MIS Quarterly* (43:2), pp. 623-647.

Lyons, K., Messinger, P. R., Niu, R. H., and Stroulia, E. 2012. "A Tale of Two Pricing Systems for Services," *Information Systems and e-Business Management* (10:1), pp. 19-42.

Maloney-Krichmar, D., and Preece, J. 2005. "A Multilevel Analysis of Sociability, Usability, and Community Dynamics in an Online Health Community," *ACM Transactions on Computer-Human Interaction (TOCHI)* (12:2), pp. 201-232.

Mäkinen, M. 2006. "Digital Empowerment as a Process for Enhancing Citizens' Participation," *E-learning and Digital Media* (3:3), pp. 381-395.

Martens, D., Provost, F., Clark, J., and de Fortuny, E. J. 2016. "Mining Massive Fine-Grained Behavior Data to Improve Predictive Analytics," *MIS Quarterly* (40:4), pp. 869-888.

Misra, R., Mukherjee, A., and Peterson, R. 2008. "Value Creation in Virtual Communities: The Case of a Healthcare Web Site," *International Journal of Pharmaceutical and Healthcare Marketing* (2:4), pp. 321-337.

Muller, E., and Doloreux, D. 2009. "What We Should Know About Knowledge-Intensive Business Services," *Technology in society* (31:1), pp. 64-72.

Müller, K. I., Alstadhaug, K. B., and Bekkelund, S. I. 2016. "Acceptability, Feasibility, and Cost of Telemedicine for Nonacute Headaches: A Randomized Study

Comparing Video and Traditional Consultations," *Journal of medical Internet research* (18:5), p. e140.

Murphy, K. P. 2012. *Machine Learning: A Probabilistic Perspective*. MIT press.

Nan, G., Wu, D., Li, M., and Tan, Y. 2018. "Optimal Freemium Strategy for Information Goods in the Presence of Piracy," *Journal of the Association for Information Systems* (19:4), pp. 266-305.

Nelson, P. 1970. "Information and Consumer Behavior," *Journal of political economy* (78:2), pp. 311-329.

Niemand, T., Mai, R., and Kraus, S. 2019. "The Zero-Price Effect in Freemium Business Models: The Moderating Effects of Free Mentality and Price–Quality Inference," *Psychology & Marketing* (early view).

Niemand, T., Tischer, S., Fritzsche, T., and Kraus, S. 2015. "The Freemium Effect: Why Consumers Perceive More Value with Free Than with Premium Offers," *36th International Conference on Information Systems*, Fort Worth.

Oestreicher-Singer, G., and Zalmanson, L. 2013. "Content or Community? A Digital Business Strategy for Content Providers in the Social Age," *MIS Quarterly* (37:2), pp. 591-616.

Ozdemir, Z. D. 2007. "Optimal Multi-Channel Delivery of Expertise: An Economic Analysis," *International Journal of Electronic Commerce* (11:3), pp. 89-105.

Park, W., Assael, H., and Chaiy, S. 1987. "Mediating Effects of Trial and Learning on Involvement-Associated Characteristics," *Journal of Consumer Marketing* (4:3), pp. 25-34.

Pawar, A., Shastri, D., and Raut, U. R. 2016. "In-Store Sampling and Impulsive Buying Behavior: An Empirical Approach," *Journal of Applied Research* (2:4), pp. 304-307.

Powers, D. M. W. 2011. "Evaluation: From Precision, Recall and F-Measure to Roc, Informedness, Markedness & Correlation," *Journal of Machine Learning Technologies* (2:1), pp. 37-63.

Raghunathan, S. 2000. "Software Editions: An Application of Segmentation Theory to the Packaged Software Market," *Journal of Management Information Systems* (17:1), pp. 87-113.

Rajala, R., and Westerlund, M. 2007. "A Business Model Perspective on Knowledge-Intensive Services in the Software Industry," *International Journal of Technoentrepreneurship* (1:1), pp. 1-20.

Rose, S., and Samouel, P. 2009. "Internal Psychological Versus External Market-Driven Determinants of the Amount of Consumer Information Search Amongst Online Shoppers," *Journal of Marketing Management* (25:1-2), pp. 171-190.

Shafer, J. C., Agrawal, R., and Mehta, M. 1996. "Sprint: A Scalable Parallel Classifier for Data Mining," *Proceedings of the 22th International Conference on Very Large Data Bases*: Morgan Kaufmann Publishers Inc., pp. 544-555.

Shampanier, K., Mazar, N., and Ariely, D. 2007. "Zero as a Special Price: The True Value of Free Products," *Marketing science* (26:6), pp. 742-757.

Shapiro, C., and Varian, H. R. 1998. "Versioning: The Smart Way to Sell Information." Boston: Harvard Business School Press, p. 106.

Shaw, S., Wherton, J., Vijayaraghavan, S., Morris, J., Bhattacharya, S., Hanson, P., Campbell-Richards, D., Ramoutar, S., Collard, A., and Hodkinson, I. 2018. "Advantages and Limitations of Virtual Online Consultations in a Nhs Acute Trust: The Vocal Mixed-Methods Study," *Health Services and Delivery Research* (6:21).

Shivendu, S., and Zhang, Z. 2015. "Versioning in the Software Industry: Heterogeneous Disutility from Underprovisioning of Functionality," *Information Systems Research* (26:4), pp. 731-753.

Silipo, R., Adae, I., Hart, A., and Berthold, M. 2015. "Seven Techniques for Data Dimensionality Reduction." Retrieved May 1, 2019, from https://www.knime.com/blog/seven-techniques-for-data-dimensionality-reduction

Singh, A. P., Joshi, H. S., Singh, A., Agarwal, M., and Kaur, P. 2018. "Online Medical Consultation: A Review," *International Journal Of Community Medicine And Public Health* (5:4), pp. 1230-1232.

Sokolova, M., and Lapalme, G. 2009. "A Systematic Analysis of Performance Measures for Classification Tasks," *Information Processing & Management* (45:4), pp. 427-437.

Song, X., Yan, X., and Li, Y. 2015. "Modelling Liking Networks in an Online Healthcare Community: An Exponential Random Graph Model Analysis Approach," *Journal of Information Science* (41:1), pp. 89-96.

Sprott, D. E., and Shimp, T. A. 2004. "Using Product Sampling to Augment the Perceived Quality of Store Brands," *Journal of Retailing* (80:4), pp. 305-315.

Statista. 2018. "Mobile App Monetization." Retrieved April 25, 2019, from https://www.statista.com/study/11480/app-monetization-statista-dossier/

Storey, C., Cankurtaran, P., Papastathopoulou, P., & Hultink, E. J. 2016. "Success factors for service innovation: A meta-analysis." Journal of Product Innovation Management, 33(5), pp. 527-548.

Sun, W., Dang, Y., and Guo, X. 2019. "The Impact of a New App Channel on Physicians' Performance: Evidence from Online Healthcare Natural Experiment," *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

Voigt, S., and Hinz, O. 2016. "Making Digital Freemium Business Models a Success: Predicting Customers' Lifetime Value Via Initial Purchase Information," *Business & Information Systems Engineering* (58:2), pp. 107-118.

Wagner, T. M., Benlian, A., and Hess, T. 2013. "The Advertising Effect of Free -- Do Free Basic Versions Promote Premium Versions within the Freemium Business Model of Music Services?," *46th Hawaii International Conference on System Sciences*: IEEE, pp. 2928-2937.

Wagner, T. M., Benlian, A., and Hess, T. 2014. "Converting Freemium Customers from Free to Premium—the Role of the Perceived Premium Fit in the Case of Music as a Service," *Electronic Markets* (24:4), pp. 259-268.

Wakefield, K. L., & Blodgett, J. G. 1999. "Customer response to intangible and tangible service factors". *Psychology & Marketing*, (16:1), pp. 51-68.

Wang, Q., Li, B., and Singh, P. V. 2018. "Copycats Vs. Original Mobile Apps: A Machine Learning Copycat-Detection Method and Empirical Analysis," *Information Systems Research* (29:2), pp. 273-291.

Wei, X. D., and Nault, B. R. 2013. "Experience Information Goods:"Version-to-Upgrade"," *Decision Support Systems* (56), pp. 494-501.

Wilson, F. 2006. "The Freemium Business Model." Retrieved May 1st, 2019, from https://avc.com/2006/03/the_freemium_bu/

Wu, H., and Deng, Z. 2019. "Knowledge Collaboration among Physicians in Online Health Communities: A Transactive Memory Perspective," *International Journal of Information Management* (49), pp. 13-33.

Wu, H., and Lu, N. 2018. "Service Provision, Pricing, and Patient Satisfaction in Online Health Communities," *International journal of medical informatics* (110), pp. 77-89.

Wu, S.-y., and Chen, P.-y. 2008. "Versioning and Piracy Control for Digital Information Goods," *Operations Research* (56:1), pp. 157-172.

Yang, H., Guo, X., Wu, T., and Ju, X. 2015. "Exploring the Effects of Patient-Generated and System-Generated Information on Patients' Online Search, Evaluation and Decision," *Electronic Commerce Research and Applications* (14:3), pp. 192-203.

Yang, H., and Zhang, X. 2019. "Investigating the Effect of Paid and Free Feedback About Physicians' Telemedicine Services on Patients' and Physicians' Behaviors: Panel Data Analysis," *Journal of medical Internet research* (21:3), p. e12156.

Yang, Y., Zhang, X., and Lee, P. K. 2019. "Improving the Effectiveness of Online Healthcare Platforms: An Empirical Study with Multi-Period Patient-Doctor Consultation Data," *International Journal of Production Economics* (207), pp. 70-80.

Yellowlees, P., Richard Chan, S., and Burke Parish, M. 2015. "The Hybrid Doctor–Patient Relationship in the Age of Technology – Telepsychiatry Consultations and the Use of Virtual Space," *International Review of Psychiatry* (27:6), pp. 476-489.

Yu, H., Xiang, K., and Yu, J. 2017. "Understanding a Moderating Effect of Physicians' Endorsement to Online Workload: An Empirical Study in Online Health-Care Communities," *2017 IEEE International Conference on Big Data (Big Data)*: IEEE, pp. 4866-4868.

Yu, L., and Liu, H. 2003. "Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution," *Proceedings of the 20th international conference on machine learning (ICML-03)*, pp. 856-863.

Zeithaml, V. A. 1988. "Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence," *Journal of marketing* (52:3), pp. 2-22.

Zhao, S., Grasmuck, S., and Martin, J. 2008. "Identity Construction on Facebook: Digital Empowerment in Anchored Relationships," *Computers in human behavior* (24:5), pp. 1816-1836.

Conclusion

We are living in an ever-digitized world where people are connected virtually, and virtualized processes are transforming how we feel, think, and act. To build an understanding of the impacts of virtualized processes in this era of digital transformation, this thesis consists of three stories that captures how technologies empower virtualized processes and generate impacts on individual behaviors, as well as how digital representations of the objects and actions in a virtualized process provide a medium to carry information and deepen collaborations among humans and machines.

Through a systematic literature review, essay 1 provides an understanding of how patients with chronic diseases digitize their body information through self-monitoring technologies. The newly created information (the digitized body and health status) is processed by patients and healthcare providers with the help of technologies, which further influences the use of technology, patient-provider interactions, their understanding of health conditions, chronic disease self-management behaviors and health outcomes. The key affordances that have been identified (i.e., preparation, data collection, user reflection and action, and social connection) bring to light the key capabilities that an ITSM system could deliver to enhance the experience of ITSM. The four intermediate outcomes (i.e., patient-provider co-management, patients learning and self-reflection, intervention satisfaction and compliance, and social interaction) show the results that an individual can achieve due to engaging in ITSM, which may serve as intermediate mechanisms to explain how ITSM generate impacts on behavior change and health outcomes. The overarching framework enables an overview of the current state of ITSM research and surfaces the gaps for future research in this area.

Essay 2 provides a meta-schema explaining how an offering's (i.e., products and services) digital representation on the platform endows it with causal capacities in which the value and quality of the offering is perceptible by human and machine agents on the platform. The newly formed cognitive frames and beliefs about the offering result in actions and interactions among human and machine agents, which further update the offering's digital representation and its causal capacity. This involves virtualized processes such as offering evaluation, transaction, relationship building and new content creation. By simultaneously considering the sense-making and action mechanisms from human actors and machines, the meta-schema contributes to our understanding of how human-machine hybrid systems on digital platforms may produce collective outcomes. It also provides a vocabulary and seven high-level mechanisms that can be leveraged by future context-specific research.

Essay 3 examines the impacts of virtualized patient-physician interactions during online medical consultation, which are enabled by various digital platform functionalities such as reviewing, consultation history display and virtual gift systems. The digital representation of consultation services may influence patients' payment decisions for high-end premium services. By adopting a machine learning approach for data mining, the results show the importance of service quality (e.g., consultation dialogue quantity, response intensity), patient source (e.g., previous offline connection), patient's online involvement (e.g., patient question quantity, social return offering) and physician reputation (e.g., title, affiliation). The study illustrates the usefulness of machine learning as a complementary approach to traditional regression-based analysis for mining massive consumer data with a high number of dimensions. The cross-validated results contribute to our understanding of the types of service that can attract payment by considering complex interactions among the service features.

Implications for Theorizing and Building Interim Theories

Besides the context-specific contributions in understanding virtualized processes, this thesis is also an exercise in building interim theories by using different reasoning approaches which open revenues for future refinement and verification. Theorizing is one of the most important missions of scientific inquiry by researchers. "Theory is King," coined in Straub's (2009) editorial for *MIS Quarterly*, set the tone for IS research to be theory-rich, and theory became the required element for publishing in good journals. In IS, we often aspire to formal and strong theories that provide systematic reasons about why events and structures do or do not occur (Sutton and Staw 1995). A good, strong theory usually includes well-defined constructs, relationships and clear boundaries, which can be evaluated by its rigor, falsifiability, generality, accuracy, simplicity, consistency, and robustness (Alter 2017). Thus, a formal theory is difficult to develop, and it is often difficult to develop a new theory using just one study. Theorizing can be an iterative process with multiple trials, modifications and verifications.

Weick (1995) conceived of these products in the early theoretical development stages as "interim theories," and he considered theory as not only a *product* but also a *process* in that a formal and strong theory often emerges over time and builds upon various interim conceptual artifacts. These interim conceptual artifacts may include lists of concepts that provide abstractions and descriptions of the phenomenon, frameworks and models that show the paths of influence and effects, propositions or hypotheses that specify the constructs and the corresponding relationships, and even patterns and arguments that provide logically integrated explanations or descriptions. Although these interim conceptual artifacts seem to be semi-products of a formal, strong theory and may not meet all the evaluation criteria of a strong theory, they are indeed important and useful because they summarize the research progress, screen the most crucial part of the phenomenon of interest, and give direction for further scientific inquiry.

Our three essays offer theoretical contributions by building interim theories. Essay 1 developed a framework to explain how ITSM may facilitate behavior change and improve health outcomes through technology use and four intermediary mechanisms. The organization of the key concepts in the framework follows the affordance actualization theory, and the relationships between the constructs are synthesized from the literature. The framework is an interim theory in that it sets the boundary of explanation by providing the key building blocks and specifies the general relations among them. However, further research is needed to formalize the statements of relationships between finer-grained constructs and verify the relationships, as well as the boundary conditions of values.

By combining the typology of social mechanism and computational mechanism perspectives, essay 2 developed a meta-schema that attempts to explain the joint action between human and machine on digital platforms. The meta-schema is proposed under clear contextual and conceptual assumptions with rigorously defined theoretical concepts, which provides a foundation for future development of middle-range theories with a specific empirical phenomenon.

Essay 3 is data-driven research with an objective to discover plausible explanations for online medical consultation payment. Simon (1977) makes distinctions between theory-driven research and data-driven research. The former approach suggests that data should come into the theory construction process in the testing stage so that researchers start from a research problem, construct the hypothesis, and are confronted with data. On the other hand, data-driven research begins with the data so that data – rather than existing theory – provides guidance for inquiry and new insight generation. With the increasing availability of massive user data, there is a recent call for integrating data-driven and theory-driven research (Elragal and Klischewski 2017; Maass et al. 2018; Rai 2016). In general, massive datasets along with appropriate analysis techniques allow for the identification of new patterns that contribute to explaining the phenomenon of interest, whereas the domain theory and its constructs and relationships generate data requirements. These two processes may happen iteratively, so that the discovery of new patterns from data may change our interpretation of the phenomenon, thus refining the application of domain theory. The domain theory may further refine data analysis (even require new data sources or data dimensions). The results of data analysis in essay 3, in combination with the literature review results, highlight service quality and patient involvement as two candidate explanations that are ranked higher than physician reputation. Future research can start from here and try to formalize the relationship between the proposed explanans and the outcome (i.e., premium payment) with additional data (e.g., service quality evaluation data, user perception data) for finer-grained analysis.

Fertile Grounds for Future Research

To build an understanding of the impacts of virtualized processes in this era of digital transformation, further research on process virtualization is needed on the interplay between technologies, digital representations, and human and machine actors. Digital representation and virtualized processes are distinct but related ideas. Whereas virtualized processes remove physical interactions among actors and objects, digital representations of the objects and actions provide a medium to carry information and generate impacts. Virtualized processes also involve actors, which currently can often be carried out by both human and machine agents. On reflection, the three essays of this thesis all touch on the relationships between process virtualization, digital representation, and actors, leading us to a deeper understanding of these three dimensions and their associated fertile grounds for future research.

The Process Virtualization Dimension

Virtualizing a process involves removing the barriers from elements such as sensory requirements (e.g., need for touch), relationship requirements (e.g., face-to-face interactions), synchronism requirements (e.g., immediate responses), and identification requirements (e.g., recognizing the qualities to distinguish a person or thing). The technologies that realize process virtualization and the virtualized process seem to exhibit a technology-process duality. On the one hand, newly available technologies and their capabilities act as external driving forces to facilitate the virtualization of processes. For example, without new wearable and insideable technologies, patients cannot easily keep track of multiple body indicators on their own. Without the support of telemedicine and secured instant messaging technologies, the self-tracked body information cannot be easily shared between patients and physicians. Without the implementation of online review and recommendation systems, consumers will be drowned by the massive amount of information, and it will be difficult to match their needs and preferences with what is offered online. IT capabilities create opportunities to virtualize a variety of processes that were previously not able to be reproduced without physical interactions. This mostly involves adapting an existing physical process to its virtual form. New workflows or functions may be added to the process due to the availability of technological functions, and further reengineering or optimization may be needed to make the virtualized process more efficient.

On the other hand, technology is constructed by actors who have specific requirements in the processes. If we take online medical consultation as an example, the e-consultation industry has emerged from the need to better match the demand and supply of high-quality healthcare resources. The platform technology has been developed from a digital information hub (i.e., providing a list of available specialists) to an asynchronous communication channel (i.e., online community for healthcare information sharing) to a type of "online shopping" site (i.e., physician's time and knowledge as digital products) and now to an hybrid online hospital which is supported by various technological functionalities such as review systems, recommendation systems, medical records management, scheduling systems, synchronous communication, social support, remote team collaboration, billing, and payment. It is the emergence of new requirements in online medical consultation that facilitates the evolution of platform technologies in this particular context.

One relevant question is how should we perceive the integration of technology and virtualized processes – is it adapting the process to new technologies, or adapting technologies to the new requirements in the processes? Technologies may be developed independently from the process being virtualized. None of the technologies or modules of the medical consultation platform are developed from virtualizing a consultation process per se, and they can be implemented in various online contexts. This may force the process virtualization to adapt to the technology – medical consultation delivery may be driven to conform to platform infrastructures that are designed for general online shopping. The further adaptation of technologies may naturally happen when the existing infrastructures and functionalities cannot sufficiently fulfill the needs of process virtualization. Future research can go beyond understanding the technology as an external enabler, and can seek to understand how process virtualization shapes new technological development, which may in turn create new opportunities to improving or even reinventing existing virtualized processes. Previous studies using structuration theory to investigate the duality of technology may help to understand the integration of technology and virtualized processes (Jones and Karsten 2008; Markus and Silver 2008; Romanow et al. 2018).

The Digital Representation Dimension

Process virtualization often requires the digital representation of the objects and actors that are involved. The availability of new technologies removes many barriers and provides new possibilities regarding the characteristics that can be represented and how to represent them in a meaningful way. For example, with the development of actigraph sensors, we are able to represent the stage and quality of sleep with numbers, indices and graphs based on body movement during sleep. Whereas the traditional blood glucose monitors using lancets and test strips can give patients with diabetes a measurement point every few hours, the new under-skin sensors can continuously display glucose levels in the cellular fluid. New technological ways of presenting information such as 3D modeling, augmented reality and virtual reality bring users new experiences and insights that they may not be able to gain though 2D pictures, videos and text.

The three essays provide examples of both static digital representations (e.g., the self-tracked body indicators, stable structure of an offering, physician reputation information) and dynamic digital representation updating processes (e.g., dynamic structure of an offering, patient involvement over time). The static digital representation may be largely constrained by the design of technology. Thus, the ontological design of the system to effectively virtualize the objects that are relevant to the process is crucial. The dynamic updating process, however, is largely influenced by actor involvement – be they the offering providers, consumers, or even a machine. Whereas such impacts from actor involvement may be conceptualized as online peer influence by many previous studies, the issue can be examined from a more sociotechnical perspective. That is, virtualizing a process not only involves pre-defining the static characteristics of an object, but also involves effectively representing the actions and interactions of various actors who contribute to the process over time. Thus, future research may put more emphasis on the dynamic aspects of digital representations and the associated impacts. This is in line with the recent development of data analytics and AI-based predictions that treat user data as fuel to constantly improve the understanding of users and provide insights to guide user actions (Forbes 2018; Ransbotham et al. 2017).

The Actor Dimension

Implementing robots and AI applications has become a trend for many online businesses. Natural language processing techniques allow chatbots to help customers navigate through sites and to interact with customers like human beings. Advanced sensor data analytics allows the delivery of context-specific and highly personalized experiences. AI-based prediction allows the discovery of new patterns and processes which may unlock new opportunities for interactions based on an individual's needs and action trajectory. It is important to explicitly consider machines (or the algorithms) as actors who participate in the virtualized processes. We increasingly see collaborations between humans and machines in surgery rooms, classrooms, organizations and virtual marketplaces (Agrawal et al. 2019; Brynjolfsson and Mcafee 2017; Kolbjørnsrud et al. 2016). The relationship between humans and machines, as well as the tasks being done, may go beyond substitution or complementarity. Collaborative decision making between humans and machines and even automated decision making by machines alone will revolutionize many aspects of our lives.

Currently, the top journals in our field of IS are calling for research on human-machine (or AI) collaboration as a hybrid system (Berente et al. 2019; Demetis and Lee 2018; Rai et al. 2019). Future research may shift from traditional IT deterministic or tool-based views of IT that focuses on technological evolution and impacts to a more integrative and social view of the human-technology ecosystem as a pervasive societal phenomenon that needs to be orchestrated and managed with a new mindset (Norman 2017; Seeber et al. 2018).

In summary, this thesis highlights the transformational impacts of virtualized processes, especially in the contexts of healthcare and digital platforms. As a starting point, the thesis takes a particular interest in individuals and the impacts of virtualized processes on individual decision making and actions. It is hoped that the three essays offer some first steps toward understanding the important role of technologies, digital representations and human-machine hybrids that influence virtual interactions and decision making among the actors. Future research can go beyond the individual level and take the directions highlighted above to investigate how technologies afford process virtualization at different stages of a business, how digital representations influence (or be influenced by) human-machine hybrids, as well as how humans and machines collaborate under different circumstances.

References

Agrawal, A., Gans, J. S., and Goldfarb, A. 2019. "Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction," *The Journal of Economic Perspectives* (33:2), pp. 31-50.

Alter, S. 2017. "Nothing Is More Practical Than a Good Conceptual Artifact... Which May Be a Theory, Framework, Model, Metaphor, Paradigm or Perhaps Some Other Abstraction," *Information Systems Journal* (27:5), pp. 671-693.

Berente, N., Gu, B., Recker, J., and Santhanam, R. 2019. "Call for Paper: MIS Quarterly Special Issue on "Managing AI"." Retrieved June 04, 2019, from https://www.misq.org/skin/frontend/default/misq/pdf/CurrentCalls/ManagingAI.pdf

Brynjolfsson, E., and Mcafee, A. 2017. "The Business of Artificial Intelligence," Harvard Business Review (July), pp. 1-20.

Demetis, D. S., and Lee, A. S. 2018. "When Humans Using the IT Artifact Becomes IT Using the Human Artifact," *Journal of the Association for Information Systems* (19:10), pp. 929-952.

Elragal, A., and Klischewski, R. 2017. "Theory-Driven or Process-Driven Prediction? Epistemological Challenges of Big Data Analytics," *Journal of Big Data* (4:1), p. 19.

Forbes. 2018. "4 Powerful Ways Artificial Intelligence Is Molding E-Commerce." Retrieved June 09, 2019, from https://www.forbes.com/sites/cognitiveworld/2018/08/05/4-powerful-ways-artificial-intelligence-is-molding-e-commerce/#7db01e524b11

Jones, M. R., and Karsten, H. 2008. "Giddens's Structuration Theory and Information Systems Research," *MIS Quarterly* (32:1), pp. 127-157.

Kolbjørnsrud, V., Amico, R., and Thomas, R. J. 2016. "How Artificial Intelligence Will Redefine Management," *Harvard Business Review* (2).

Maass, W., Parsons, J., Purao, S., Storey, V. C., and Woo, C. 2018. "Data-Driven Meets Theory-Driven Research in the Era of Big Data: Opportunities and Challenges for Information Systems Research," *Journal of the Association for Information Systems* (19:12), pp. 1253-1273.

Markus, M. L., and Silver, M. S. 2008. "A Foundation for the Study of IT Effects: A New Look at Desanctis and Poole's Concepts of Structural Features and Spirit," *Journal of the Association for Information Systems* (9:10), pp. 609-632.

Norman, D. 2017. "Design, Business Models, and Human-Technology Teamwork: As Automation and Artificial Intelligence Technologies Develop, We Need to Think Less About Human-Machine Interfaces and More About Human-Machine Teamwork," *Research-Technology Management* (60:1), p. 26.

Rai, A. 2016. "Editor's Comments: Synergies between Big Data and Theory," *MIS Quarterly* (40:2), pp. iii-ix.

Rai, A., Constantinides, P., and Sarker, S. 2019. "Editor's Comments: Next-Generation Digital Platforms: Toward Human–AI Hybrids," *Management Information Systems Quarterly* (43:1), pp. iii-ix.

Ransbotham, S., Kiron, D., Gerbert, P., and Reeves, M. 2017. "Reshaping Business with Artificial Intelligence: Closing the Gap between Ambition and Action," *MIT Sloan Management Review* (59:1), pp.11-13.

Romanow, D., Rai, A., and Keil, M. 2018. "Cpoe-Enabled Coordination: Appropriation for Deep Structure Use and Impacts on Patient Outcomes," *MIS Quarterly* (42:1), pp. 189-A111.

Seeber, I., Bittner, E., Briggs, R. O., de Vreede, G.-J., de Vreede, T., Druckenmiller, D., Maier, R., Merz, A. B., Oeste-Reiß, S., and Randrup, N. 2018. "Machines as Teammates: A Collaboration Research Agenda," *Proceedings of the 51st Hawaii International Conference on System Sciences*.

Simon, H. A. 1977. *Models of Discovery : And Other Topics in the Methods of Science*. Dordrecht, Holland ;: D. Reidel Pub. Co.

Straub, D. W. 2009. "Editor's Comments: Why Top Journals Accept Your Paper," *MIS Quarterly* (33:3), pp. III-X.

Sutton, R. I., and Staw, B. M. 1995. "What Theory Is Not," *Administrative Science Quarterly* (40:3), pp. 371-384.

Weick, K. E. 1995. "What Theory Is Not, Theorizing Is," *Administrative Science Quarterly* (40:3), pp. 385-390.

Appendix A – Coding Results for Essay 1

Table A.1 Profile of the Studies by IT and Disease Type								
	Mobile/tablet app	**Website**	**Medical Device**	**Wearable**	**IVR**	**PDA**	**PC software**	**Other**
Obesity	Allen et al. (2013), Carter et al. (2013), Cushing et al. (2011), Hutchesson et al. (2015), Kolodziejczyk et al.(2014), Laing et al. (2015), Partridge et al. (2016), Tsai et al. (2007), Turner-McGrievy et al. (2013), Wharton et al. (2014), Chen et al. (2017), Hales et al. (2017), Jospe et al. (2017a), Mummah et al. (2017), Sasai et al. (2017), Spring et al. (2017), Turner-McGrievy et al. (2017)	Carter et al. (2013), Dennison et al. (2014), Hutchesson et al. (2015), Krukowski et al. (2013), Morgan et al. (2014), Partridge et al. (2016), Ruotsalainen et al. (2015), Shuger et al. (2011), Thomas et al. (2015), Webber et al. (2010), Wolin et al. (2015), Carels et al. (2017), Fuller et al. (2017), Jakicic et al. (2016), Painter et L. (2017), Rader et al. (2017), Tu et al. (2017)	Donaldson and Normand (2009), Jospe et al. (2017b), Moho Shaiful et al. (2017), Painter et L. (2017)	Cadmus-Bertram et al. (2015), Shuger et al. (2011), Carels et al. (2017), Jakicic et al. (2016), Moho Shaiful et al. (2017), Painter et L. (2017), Sasai et al. (2017), Turner-McGrievy et al. (2017), Cadmus-Bertram et al. (2013), Morgan et al. (2014), Nicklas et al. (2014), Ruotsalainen et al. (2015), Steinberg et al. (2013), Aguiar et al. (2017), Spring et al. (2017), Tu et al. (2017)	Steinberg et al. (2014), Wolin et al. (2015), Steinberg et al. (2017)	Acharya et al. (2011), Ambeba et al. (2015), Burke et al. (2012), Conroy et al. (2011), Turk et al. (2013), Wang et al. (2012), Yon et al. (2006)	Chambliss et al. (2011), Schroder (2011)	Chung et al. (2015), Chung et al. (2016), Williamson et al. (2010), Sidhu et al. (2016), Sasai et al. (2017), Tu et al. (2017)
Diabetes	Or and Tao (2016), Roblin (2011), Storni (2014), Storni (2014b), di Bartolo et al. (2017), Garg et al. (2017), Gu et al. (2017), Hansen et al. (2017), Irace et al. (2017), Munster-Segev et al. (2017), Piras and Miele (2017), Sieber et al. (2017)	Glasgow et al. (2011), Greenwood et al. (2015), Hinnen et al. (2015), Raiff and Dallery (2010), Caballero-Ruiz et al. (2017), Iljaz et al. (2017), Irace et al. (2017), Kempf et al. (2017)	Greenwood et al. (2015), Hinnen et al. (2015), Or and Tao (2016), Raiff and Dallery (2010), Roblin (2011), Sevick et al. (2008), Caballero-Ruiz et al. (2017), Cosson et al. (2017), di Bartolo et al. (2017), Downing et al. (2017), Garg et al. (2017), Goffinet et al. (2017), Haak et al. (2017), Irace et al. (2017), Ji et al. (2017), Kempf et al. (2017), Lee et al. (2017), Mathieu-Fritz et al. (2017), Nishimura et al. (2017), Olafsdottir et al. (2017), Paula et al. (2017), Polonsky et al. (2017), Selvan et al. (2017), Sieber et al.	Edge et al. (2017), Glasgow et al. (2011), Biddle et al. (2017), Kempf et al. (2017)	Glasgow et al. (2011)	Sevick et al. (2010), Sevick et al. (2008)	Paula et al. (2017)	Vaughn-Cooke et al. (2015)

Table A.1 Profile of the Studies by IT and Disease Type

	Mobile/tablet app	Website	Medical Device	Wearable	IVR	PDA	PC software	Other
			(2017), Young et al. (2017)					
Psychiatric	Faurholt-Jepsen et al. (2015a), Faurholt-Jepsen et al. (2015b), Festersen and Corradini (2014), Scharer et al. (2015), Tregarthen et al. (2015), Tsanas et al. (2016), Abrantes et al. (2017), Boyd et al. (2017), Mantani et al. (2017)	Jones et al. (2014), Nørregaard et al. (2014), Tsanas et al. (2016)	Simons et al. (2017)	Abrantes et al. (2017), Boyd et al. (2017)	\	\	Bauer et al. (2009)	Murnane et al. (2016), Matthews et al. (2017a), McKnight et al. (2017)
Cardiac	Karhula et al. (2015)	Dorsch et al. (2015)	Karhula et al. (2015), Andersen et al. (2017)	Izawa et al. (2006), Vogel et al. (2017)	\	\	\	Coppini et al. (2017)
Cancer	Timmerman et al. (2016), Mouzouras et al. (2017)	Berry et al. (2015)	Timmerman et al. (2016)	Gell et al. (2017)	\	\	\	Hall and Murchie (2014), Hermansen-Kobulnicky and Purtzer (2014)
Nerve-related	Ayobi et al. (2017)	Jongen et al. (2015), Barakat et al. (2017)	\	Ayobi et al. (2017), Mentis et al. (2017)	\	\	Ayobi et al. (2017)	Ayobi et al. (2017)
HIV	Swendeman et al. (2015)	Swendeman et al. (2015)	\	Aharonovich et al. (2017b)	Aharonovich et al. (2006), Aharonovich et al. (2017a)	\	\	\
Hypertension	Kendall et al. (2015), Or and Tao (2016)	Wolin et al. (2015)	Nakano et al. (2011), Or and Tao (2016)	\	Wolin et al. (2015)	\	Nakano et al. (2011)	Storni (2010)
Other	Fukuoka et al. (2011), Ramanathan et al. (2013), Adams et al. (2017), Plow and Golding (2017), Ryan et al. (2012), Welch et al. (2013), Cai et al. (2017), Dietrich et al. (2017), Eikey et al. (2017), Hostler et al. (2017), Isetta et al. (2017), Sage et al. (2017), Velardo et al. (2017), Zhu et al. (2017)	Langstrup and Winthereik (2008), Felipe et al. (2015), Johnston et al. (2009), Ma et al. (2013), Pedersen et al. (2012), Umapathy et al. (2015), Dietrich et al. (2017), Hostler et al. (2017), McDonald et al. (2017)	Grönvall and Verdezoto (2013), Velardo et al. (2017)	Felipe et al. (2015), Goto et al. (2014)	Naylor et al. (2008)	Dowell and Welch (2006), Stark et al. (2011), Welch et al. (2007)	Welch et al. (2007)	Bonilla et al. (2015), Verdezoto and Gronvall (2016), Chung et al. (2015), Chung et al. (2016)

Table A.2 Effects of ITSM Affordance Bundles on Chronic Care Goal Achievement							
ITSM Functionality Combinations	**Outcome: Goal Achievement**						
	Physical activity	**Diet**	**Other Behavior Change**	**Weight**	**Quality of life**	**Symptom**	**Medication**
Auto, Display	Vogel et al. (2017)	Kempf et al. (2017)	Boyd et al. (2017)	Carels et al. (2017), Kempf et al. (2017), Nishimura et al. (2017)	Kempf et al. (2017) *[di Bartolo et al. (2017)]*	di Bartolo et al. (2017), Haak et al. (2017), Ji et al. (2017), Kempf et al. (2017), Nishimura et al. (2017) *[Goffinet et al. (2017)]*	Kempf et al. (2017)
Manual, Auto, Display	Donaldson and Normand (2009) *[Goto et al. (2014)]*	Donaldson and Normand (2009) *[Welch et al. (2013)]*	\	Karhula et al. (2015), Shuger et al. (2011), Welch et al. (2013), Shaiful et al. (2017) *[Shaiful et al. (2017)]*	*[Karhula et al. (2015)]*	Goto et al. (2014), Karhula et al. (2015), Downing et al. (2017), Shaiful et al. (2017), Sieber et al. (2017)	\
Manual, Push	Wang et al. (2012)	Wang et al. (2012)	\	Sidhu et al. (2016), Thomas et al. (2015), Turk et al. (2013) *[Wang et al. (2012)]*	\	Naylor et al. (2008) *[Simons et al. (2017)]*	Naylor et al. (2008)
Manual, Display, Push	Conroy et al. (2011)	Ambeba et al. (2015)	Swendeman et al. (2015)	Burke et al. (2012), Wharton et al. (2014)	\	\	\
Manual, Display	\	Acharya et al. (2011), Schroder (2011) *[Schroder (2011)]*	*[Aharonovich et al. (2006)]*	Acharya et al. (2011), Schroder (2011), Aguiar et al. (2017)	\	\	\
Manual, Display, Push, Patient-Provider	\	\	\	\	\	*[Faurholt-Jepsen et al. (2015), Pedersen et al. (2012)]*	*[Pedersen et al. (2012)]*
Auto, Display, Push	*[Biddle et al. (2017)]*	\	\	Munster-Segev et al. (2017)	\	Munster-Segev et al. (2017)	\
Manual, Auto, Push	Morgan et al. (2014)	Morgan et al. (2014)	\	Morgan et al. (2014), Steinberg et al. (2013)	\	\	\
Goal, Manual, Display	*[Allen et al. (2013); Jospe et al. (2017a)]*	*[Allen et al. (2013); Jospe et al. (2017a)]*	\	*[Allen et al. (2013); Jospe et al. (2017a)]*	\	*[Jospe et al. (2017a)]*	\

* There are 43 different combinations of ITSM functionality among the studies in theme 3. In this table, we only list the combinations used in at least two studies.

** IT functionality: Goal - goal; Manual - manual entry; Auto - auto capture; Display - data display; Push - push message; Patient-Pro - patient-provider connection. The other IT functionalities (education, gamification, and peer-to-peer interaction) do not appear in this table as they were not among any of the combinations used in at least two studies.

*** Studies in italicized brackets have non-supportive or mixed results.

Table A.3 Non-IT Components and Chronic Care Goal Achievement							
Non-IT components combinations	**Outcome: Goal Achievement**						
	Physical activity	**Diet**	**Other Behavior Change**	**Weight**	**Quality of life**	**Symptom**	**Medication**
None	Izawa et al. (2006), Ruotsalainen et al. (2015), Gell et al. (2017) *[Jones et al. (2014), Laing et al. (2015), Ruotsalainen et al. (2015)]*	Turner-McGrievy et al. (2013), Mummah et al. (2017) *[Jones et al. (2014), Laing et al. (2015), Welch et al. (2013)]*	Swendeman et al. (2015)	Sidhu et al. (2016), Turner-McGrievy et al. (2013), Welch et al. (2013) *[Jones et al. (2014), Laing et al. (2015), Ruotsalainen et al. (2015)]*	Dorsch et al. (2015)	Or and Tao (2016), Dietrich et al. (2017), Gell et al. (2017), Mantani et al. (2017) *[Faurholt-Jepsen et al. (2015), Laing et al. (2015), Or and Tao (2016), Umapathy et al. (2015)]*	*[Dietrich et al. (2017)]*
Education only	Cadmus-Bertram et al. (2013), Goto et al. (2014), Vogel et al. (2017) *[Goto et al. (2014), Biddle et al. (2017)]*	*[Schroder (2011)]*	*[Aharonovich et al. (2006)]*	Cadmus-Bertram et al. (2013) Carter et al. (2013) Schroder (2011), Shuger et al. (2011), Steinberg et al. (2013), Aguiar et al. (2017), Nishimura et al. (2017)	di Bartolo et al. (2017)	Goto et al. (2014), di Bartolo et al. (2017), Garg et al. (2017), Nishimura et al. (2017), Sieber et al. (2017) *[Pedersen et al. (2012), Garg et al. (2017), Goffinet et al. (2017)]*	*[Pedersen et al. (2012)]*
Goal only	Turner-McGrievy et al. (2017)	*[Turner-McGrievy et al. (2017)]*	\	Wharton et al. (2014), Turner-McGrievy et al. (2017)	\	\	\
Feedback only	\	Barakat et al. (2017)	\	Cadmus-Bertram et al. (2013)	*[Polonsky et al. (2017)]*	Haak et al. (2017), Iljaz et al. (2017)	\
Education + Goal	Donaldson and Normand (2009), Morgan et al. (2014), Wang et al. (2012) *[Allen et al. (2013), Jospe et al. (2017)]*	Ambeba et al. (2015), Donaldson and Normand (2009), Morgan et al. (2014), Wang et al. (2012) *[Allen et al. (2013), Jospe et al. (2017)]*	*[Aharonovich et al. (2017b)]*	Karhula et al. (2015), Morgan et al. (2014), Thomas et al. (2015) *[Allen et al. (2013), Wang et al. (2012), Jospe et al. (2017a)]*	*[Karhula et al. (2015)]*	Karhula et al. (2015), Downing et al. (2017), Jospe et al. (2017a) *[Hansen et al. (2017), Jospe et al. (2017a)]*	Aharonovich et al. (2017b)

Table A.3 Non-IT Components and Chronic Care Goal Achievement							
Non-IT components combinations	**Outcome: Goal Achievement**						
	Physical activity	**Diet**	**Other Behavior Change**	**Weight**	**Quality of life**	**Symptom**	**Medication**
Education + Feedback	\	Acharya et al. (2011), Kempf et al. (2017) *[Dowell and Welch (2006)]*	\	Acharya et al. (2011), Kempf et al. (2017), Munster-Segev et al. (2017)	Kempf et al. (2017) *[Young et al. (2017)]*	Ji et al. (2017), Kempf et al. (2017), Munster-Segev et al. (2017) *[Simons et al. (2017), Young et al. (2017)]*	Kempf et al. (2017)
Education + Social	\	\	\	Shaiful et al. (2017)	\	Shaiful et al. (2017)	\
Education + Goal + Feedback	Cadmus-Bertram et al. (2015), Conroy et al. (2011), Nicklas et al. (2014), Abrantes et al. (2017), Jakicic et al. (2016), Plow and Golding (2017) *[Abrantes et al. (2017), Jakicic et al. (2016), Plow and Golding (2017)]*	Nicklas et al. (2014) *[Jakicic et al. (2016)]*	*[Abrantes et al. (2017)]*	Burke et al. (2012), Chambliss et al. (2011), Nicklas et al. (2014), Turk et al. (2013), Carels et al. (2017), Jakicic et al. (2016) *[Abrantes et al. (2017), Jakicic et al. (2016)]*	Ryan et al. (2012)	Chambliss et al. (2011), Abrantes et al. (2017), Steinberg et al. (2017) *[Greenwood et al. (2015), Ryan et al. (2012), Abrantes et al. (2017), Jakicic et al. (2016)]*	Aharonovich et al. (2017a) *[Plow and Golding (2017)]*
Education + Feedback + Social	\	\	\	\	\	Naylor et al. (2008)	Naylor et al. (2008)
Education + Goal + Feedback + Social	*[Sasai et al. (2017)]*	\	\	Spring et al. (2017)	\	Sasai et al. (2017)	\

** Studies in italicized brackets have non-supportive or mixed results.

Table A.4 Key IT Functionalities that Enable ITSM Affordances			
IT Functionality	Studies that included this functionality		
	2006-2009	2010-2013	2014-2017
Preparation Affordance			
Deliver educational content (IT-delivered content for increasing knowledge of the device, the disease or of its self-management)	N/A	Cadmus-Bertram et al. (2013), Carter et al. (2013), Chambliss et al. (2011), Glasgow et al. (2011), Krukowski et al. (2013), Webber et al. (2010)	Dennison et al. (2014), Dorsch et al. (2015), Festersen and Corradini (2014), Greenwood et al. (2015), Hinnen et al. (2015), Kolodziejczyk et al.(2014), Or and Tao (2016), Partridge et al. (2016), Timmerman et al. (2016), Umapathy et al. (2015), Wolin et al. (2015), Aharonovich et al. (2017b), Barakat et al. (2017), Cai et al. (2017), Coppini et al. (2017), Dietrich et al. (2017), Hales et al. (2017), Hostler et al. (2017), Iljaz et al. (2017), Isetta et al. (2017), Jakicic et al. (2016), Lee et al. (2017), Mantani et al. (2017), Mouzouras et al. (2017), Mummah et al. (2017), Rader et al. (2017), Sage et al. (2017), Sasai et al. (2017), Tu et al. (2017), Turner-McGrievy et al. (2017), Velardo et al. (2017), Young et al. (2017)
Goal setting (IT suggests or assigns goals or allows users to set and modify their own goals)	Johnston et al. (2009)	Allen et al. (2013), Cadmus-Bertram et al. (2013), Carter et al. (2013), Chambliss et al. (2011), Sevick et al. (2010), Stark et al. (2011)	Cadmus-Bertram et al. (2015), Dennison et al. (2014), Steinberg et al. (2014), Abrantes et al. (2017), Eikey et al. (2017), Hales et al. (2017), Hostler et al. (2017), Jospe et al. (2017a), Mummah et al. (2017), Painter et L. (2017), Plow and Golding (2017), Steinberg et al. (2017), Tu et al. (2017)
Data Collection Affordance			
Data entry interface (User-initiated SM data entry. Can offer different levels of flexibility of input such as guided response or open entry)	Aharonovich et al. (2006), Donaldson and Normand (2009), Dowell and Welch (2006), Johnston et al. (2009), Naylor et al. (2008), Sevick et al. (2008), Tsai et al. (2007), Welch et al. (2007)	Acharya et al. (2011), Allen et al. (2013), Burke et al. (2012), Carter et al. (2013), Chambliss et al. (2011), Conroy et al. (2011), Cushing et al. (2011), Glasgow et al. (2011), Krukowski et al. (2013), Ma et al. (2013), Pedersen et al. (2012), Raiff and Dallery (2010), Roblin (2011), Ryan et al. (2012), Schroder (2011), Sevick et al. (2010), Shuger et al. (2011), Stark et al. (2011), Steinberg et al. (2013), Turk et al. (2013), Turner-McGrievy et al. (2013), Wang et al. (2012), Webber et al. (2010), Welch et al. (2013), Williamson et al. (2010)	Ambeba et al. (2015), Berry et al. (2015), Bonilla et al. (2015), Dennison et al. (2014), Dorsch et al. (2015), Faurholt-Jepsen et al. (2015), Festersen and Corradini (2014), Goto et al. (2014), Greenwood et al. (2015), Hutchesson et al. (2015), Jones et al. (2014), Jongen et al. (2015), Karhula et al. (2015), Kendall et al. (2015), Kolodziejczyk et al. (2014), Laing et al. (2015), Morgan et al. (2014), Nicklas et al. (2014), Partridge et al. (2016), Ruotsalainen et al. (2015), Sidhu et al. (2016), Steinberg et al. (2014), Storni (2014), Swendeman et al. (2015), Thomas et al. (2015), Tsanas et al. (2016), Umapathy et al. (2015), Wharton et al. (2014), Wolin et al. (2015), Adams et al. (2017), Aguiar et al. (2017), Aharonovich et al. (2017a), Aharonovich et al. (2017b), Ayobi et al. (2017), Barakat et al. (2017), Caballero-Ruiz et al. (2017), Dietrich et al. (2017), Downing et al. (2017), Eikey et al. (2017), Fuller et al. (2017), Gu et al. (2017), Hales et al. (2017), Hansen et al. (2017), Hostler et al. (2017), Iljaz et al. (2017), Isetta et al. (2017), Jospe et al. (2017a), Jakicic et al. (2016), Lee et al. (2017), Mantani et al. (2017), McDonald et al. (2017), McKnight et al. (2017), Moho Shaiful et al. (2017), Mouzouras et al. (2017), Mummah et al. (2017), Painter et L. (2017), Plow and Golding (2017), Rader et al. (2017), Sage et al. (2017), Selvan et al. (2017), Sieber et al. (2017), Simons et al. (2017), Spring et al. (2017), Steinberg et al. (2017), Tu et al. (2017), Turner-McGrievy et al. (2017), Velardo et al. (2017)

Table A.4 Key IT Functionalities that Enable ITSM Affordances			
IT Functionality	Studies that included this functionality		
	2006-2009	2010-2013	2014-2017
Auto capture (Automatic measuring of SM efforts)	Donaldson and Normand (2009), Sevick et al. (2008)	Cadmus-Bertram et al. (2013), Carter et al. (2013), Nakano et al. (2011), Raiff and Dallery (2010), Roblin (2011), Ryan et al. (2012), Shuger et al. (2011), Steinberg et al. (2013), Welch et al. (2013)	Cadmus-Bertram et al. (2015), Felipe et al. (2015), Goto et al. (2014), Greenwood et al. (2015), Hinnen et al. (2015), Karhula et al. (2015), Kolodziejczyk et al.(2014), Laing et al. (2015), Morgan et al. (2014), Nicklas et al. (2014), Or and Tao (2016), Partridge et al. (2016), Ruotsalainen et al. (2015), Timmerman et al. (2016), Abrantes et al. (2017), Andersen et al. (2017), Ayobi et al. (2017), Biddle et al. (2017), Boyd et al. (2017), Caballero-Ruiz et al. (2017), Carels et al. (2017), Coppini et al. (2017), Cosson et al. (2017), di Bartolo et al. (2017), Downing et al. (2017), Edge et al. (2017), Garg et al. (2017), Gell et al. (2017), Goffinet et al. (2017), Gu et al. (2017), Haak et al. (2017), Irace et al. (2017), Ji et al. (2017), Jospe et al. (2017b), Kempf et al. (2017), Jakicic et al. (2016), Lee et al. (2017), Mathieu-Fritz et al. (2017), Mentis et al. (2017), Moho Shaiful et al. (2017), Munster-Segev et al. (2017), Nishimura et al. (2017), Olafsdottir et al. (2017), Painter et L. (2017), Paula et al. (2017), Piras and Miele (2017), Plow and Golding (2017), Polonsky et al. (2017), Sasai et al. (2017), Selvan et al. (2017), Sieber et al. (2017), Spring et al. (2017), Tu et al. (2017), Turner-McGrievy et al. (2017), Velardo et al. (2017), Vogel et al. (2017), Young et al. (2017)
User Reflection and Action Affordance			
Data display (IT offers graphical, numerical, or textual feedback of the SM results with (1) raw data (2) simple aggregation, and/or (3) evaluative information that relates the data to a target, goal or threshold)	Aharonovich et al. (2006), Donaldson and Normand (2009), Johnston et al. (2009), Sevick et al. (2008), Tsai et al. (2007), Welch et al. (2007)	Acharya et al. (2011), Allen et al. (2013), Burke et al. (2012), Cadmus-Bertram et al. (2013), Carter et al. (2013), Chambliss et al. (2011), Conroy et al. (2011), Cushing et al. (2011), Glasgow et al. (2011), Nakano et al. (2011), Pedersen et al. (2012), Raiff and Dallery (2010), Roblin (2011), Ryan et al. (2012), Schroder (2011), Sevick et al. (2010), Shuger et al. (2011), Stark et al. (2011), Webber et al. (2010), Welch et al. (2013), Williamson et al. (2010)	Ambeba et al. (2015), Berry et al. (2015), Bonilla et al. (2015), Cadmus-Bertram et al. (2015), Dennison et al. (2014), Dorsch et al. (2015), Faurholt-Jepsen et al. (2015), Felipe et al. (2015), Festersen and Corradini (2014), Goto et al. (2014), Greenwood et al. (2015), Hinnen et al. (2015), Hutchesson et al. (2015), Jones et al. (2014), Jongen et al. (2015), Karhula et al. (2015), Kendall et al. (2015), Laing et al. (2015), Or and Tao (2016), Partridge et al. (2016), Ruotsalainen et al. (2015), Steinberg et al. (2014), Storni (2014), Swendeman et al. (2015), Timmerman et al. (2016), Tsanas et al. (2016), Umapathy et al. (2015), Wharton et al. (2014), Wolin et al. (2015), Abrantes et al. (2017), Adams et al. (2017), Aguiar et al. (2017), Aharonovich et al. (2017b), Andersen et al. (2017), Ayobi et al. (2017), Biddle et al. (2017), Boyd et al. (2017), Caballero-Ruiz et al. (2017), Cai et al. (2017),Carels et al. (2017), Coppini et al. (2017), Cosson et al. (2017), di Bartolo et al. (2017), Downing et al. (2017), Edge et al. (2017), Eikey et al. (2017), Fuller et al. (2017), Garg et al. (2017), Goffinet et al. (2017), Gu et al. (2017), Haak et al. (2017), Hales et al. (2017), Hansen et al. (2017), Hostler et al. (2017), Irace et al. (2017), Isetta et al. (2017), Ji et al. (2017), Jospe et al. (2017a), Jospe et al. (2017b), Kempf et al. (2017), Jakicic et al. (2016), Mantani et al. (2017), Mathieu-Fritz et al. (2017), McDonald et al. (2017), McKnight et al. (2017), Mentis et al. (2017), Moho Shaiful et al. (2017), Mouzouras et al. (2017), Mummah et al. (2017), Munster-Segev et al. (2017), Nishimura et al. (2017), Olafsdottir et al. (2017), Painter et L. (2017), Paula et al. (2017),Piras and Miele (2017),Plow and Golding (2017),Rader et al. (2017),Sage et al. (2017),Sasai et al. (2017),Selvan et al. (2017),Sieber et al. (2017),Spring et al. (2017), Steinberg et al. (2017),Tu

Table A.4 Key IT Functionalities that Enable ITSM Affordances			
IT Functionality	Studies that included this functionality		
	2006-2009	2010-2013	2014-2017
			et al. (2017), Turner-McGrievy et al. (2017),Velardo et al. (2017), Vogel et al. (2017),Young et al. (2017)
Push messages (IT delivers messages or prompts which can be 1) pre-set based on user preference or schedule, or 2) data-driven by users' own SM data)	Naylor et al. (2008); Tsai et al. (2007)	Burke et al. (2012), Carter et al. (2013), Chambliss et al. (2011), Conroy et al. (2011), Cushing et al. (2011), Glasgow et al. (2011), Nakano et al. (2011), Pedersen et al. (2012), Steinberg et al. (2013), Turk et al. (2013), Wang et al. (2012)	Ambeba et al. (2015), Berry et al. (2015), Dennison et al. (2014), Faurholt-Jepsen et al. (2015), Greenwood et al. (2015), Jones et al. (2014), Kendall et al. (2015), Morgan et al. (2014), Partridge et al. (2016), Sidhu et al. (2016), Steinberg et al. (2014), Swendeman et al. (2015), Thomas et al. (2015), Tsanas et al. (2016), Umapathy et al. (2015), Wharton et al. (2014), Wolin et al. (2015), Aharonovich et al. (2017b), Barakat et al. (2017), Biddle et al. (2017), Caballero-Ruiz et al. (2017), Cai et al. (2017), Coppini et al. (2017), Fuller et al. (2017), Hales et al. (2017), Hostler et al. (2017), Iljaz et al. (2017), Irace et al. (2017), Isetta et al. (2017), Jakicic et al. (2016), Mantani et al. (2017), McDonald et al. (2017), Mouzouras et al. (2017), Mummah et al. (2017), Munster-Segev et al. (2017), Piras and Miele (2017), Sage et al. (2017), Simons et al. (2017), Steinberg et al. (2017), Tu et al. (2017), Velardo et al. (2017), Young et al. (2017)
Gamification (Gamify ITSM tasks or SM results display)	N/A	Glasgow et al. (2011), Raiff and Dallery (2010)	Jones et al. (2014), Abrantes et al. (2017), Adams et al. (2017), Ayobi et al. (2017), Cai et al. (2017), Dietrich et al. (2017), Eikey et al. (2017), Hales et al. (2017), Hostler et al. (2017), Mantani et al. (2017), Mummah et al. (2017), Sage et al. (2017), Sasai et al. (2017), Tu et al. (2017)
Social Connections Affordance			
Patient-provider connection (IT-mediated patient-provider communication and collaboration)	N/A	Chambliss et al. (2011), Ma et al. (2013), Pedersen et al. (2012), Ryan et al. (2012), Webber et al. (2010)	Faurholt-Jepsen et al. (2015), Festersen and Corradini (2014), Greenwood et al. (2015), Jongen et al. (2015), Ruotsalainen et al. (2015), Timmerman et al. (2016), Umapathy et al. (2015), Caballero-Ruiz et al. (2017), Cai et al. (2017), Garg et al. (2017), Hansen et al. (2017), Iljaz et al. (2017), Irace et al. (2017), Mantani et al. (2017), Mouzouras et al. (2017), Painter et L. (2017), Piras and Miele (2017), Rader et al. (2017), Velardo et al. (2017)
Peer-to-peer interaction (IT-mediated social features that allow social comparison or peer-to-peer interaction)	N/A	Allen et al. (2013), Cadmus-Bertram et al. (2013), Carter et al. (2013), Glasgow et al. (2011), Krukowski et al. (2013), Turner-McGrievy et al. (2013)	Jones et al. (2014), Kolodziejczyk et al. (2014), Laing et al. (2015), Partridge et al. (2016), Ruotsalainen et al. (2015), Cai et al. (2017), Dietrich et al. (2017), Eikey et al. (2017), Hales et al. (2017), Mummah et al. (2017), Rader et al. (2017), Spring et al. (2017), Tu et al. (2017)

Appendix B – Literature Review Results for Essay 2

Table B1. Study Profile by Outlets and Publication Year (N=39)

Journal	Year												Total
	2004	2005	2006	2007	2008	2012	2013	2014	2015	2016	2017	2018	
EJIS							1						1
ISJ											1	3	4
ISR	1	1	1		2	1	2		1	1	1		11
JAIS			1										1
JIT				1							2		3
JMIS						1		3		1	2	3	10
MISQ					1	1		1		1	3	2	9
Total	**1**	**1**	**2**	**1**	**3**	**3**	**3**	**4**	**1**	**3**	**9**	**8**	**39**

Table B2. Study Profile by Study Characteristics (N=39)

Platform	Year												Total
	2004	2005	2006	2007	2008	2012	2013	2014	2015	2016	2017	2018	
crowdfunding							1			2	1	4	8
group-buying								1				1	2
online auction	1	1			1	1		1		1			6
online knowledge market				1							1		2
online retail			2		2	2	2	1			4	1	14
other								1	1		1		3
P2P lending											2	2	4
Total	**1**	**1**	**2**	**1**	**3**	**3**	**3**	**4**	**1**	**3**	**9**	**8**	**39**

Table B3. Detailed Coding of the Studies by Empirical Model and Offered Explanation

Study	Empirical Model				Offered explanation & corresponding mechanism
	Offering causal capacity	Agent cognitive frames	Agent actions	Collective outcome	
Benlian et al. (2012)	Online product recommendation (provider recommendation vs. consumer reviews); Product type (moderator)	Trusting beliefs; perceived usefulness; perceived ease of use; perceived affective quality	Intention to purchase	n/a	Product recommendation → user evaluation → purchase intention [M1a, M2a]
Burtch et al. (2013)	Others' contribution frequency; previous contribution; page views; remaining budget, search trends	n/a	Crowdfunding contribution	n/a	Others' prior decision → reinforcement (reciprocity, fairness, social norms) or substitute (altruism) → current contribution [M1a, M2a]
Burtch et al. (2016)	Contributor information embedded in previous contributions	n/a	Sbsequent contribution	n/a	Absence of information → greater uncertainty & lower confidence → indecision& inaction [M1a, M2a]
Burtch et al. (2018)	Point provision (yes or no); number of days remaining; fundraising progress		Campaign contribution (yes or no; amount)		Point provision →visitors' sensitivity to prior capital accumulation (observational learning); signal of venture quality; eliminate concerns about partial provision → contribution [M1a, M2a]
Carmi et al. (2017)	Exogenous shock (Oprah Winfrey TV show; New York Times); pre-event average sales rank; book network distance (to the reviewed book)	n/a	n/a	Sales rank	Other (external factor)
Chen et al. (2015)	Previous sales rank; previous bulletins; previous friend updates; price; release	n/a	n/a	Sales rank	n/a
Feller et al. (2017)	Explicit information expectation (hard financial data); Implicit information expectation (soft financial	n/a	Likelihood of repaying investment; likelihood of receiving investment	n/a	Information sharing → show social identity and personal cues → reliability of borrowers → lending [M1a, M2a]

Study	Empirical Model: Offering causal capacity	Agent cognitive frames	Agent actions	Collective outcome	Offered explanation & corresponding mechanism
	data; humanising data; appeals)				
Forman et al. (2008)	Reviewer identity description; shared location	n/a	n/a	sales	Review & review source → heuristic shaping of product evaluation [M1a]
Ge et al. (2017)	Borrower's disclosure of social media account; borrower's social network size; social media message posting	n/a	Default the loan	n/a	Borrower's social media →positive social image & establish positive reputation → avoid loan default to reduce social stigma cost [M1a, M2a]
Ghose et al. (2006)	Amazon price; used price	n/a	n/a	Sales rank	n/a
Gleasure et al. (2017)	Material aspects of crowdfunding technology (reviews, video, synopsis, etc.); Material aspects of internally managed production activities (editings, uploading, publisher, etc.)	n/a	Contribution practice (passive vs. active)	n/a	Enacted material aspects of crowdfunding technology and production activities → social drivers (willingness, desires, social identity) → contribution practices [M1a, M2a]
Gregg and Walczak (2008)	E-image (profesionalism of business identity, listing style); new vs. used product (moderator)	n/a	Willingness to transact	Price premium	E-image+ product type → confidence in product or company quality → willingness to transact e-image+ product type→ reducing transaction risk→ price premium [M1a, M5]
Gu et al. (2012)	External WOM vs. internal WOM (rating, number of review); price; search interest	n/a	n/a	Sales rank	External WOM → reduce perceived risk → involvement & information search behavior [M1a, M2a]
Guo et al. (2017)	Status capital; decisional capital; doctor group	n/a	n/a	Social return; economic return	Professional capital (status & decisional) → trust & perceptions on doctors' ability and willingness → provide social and economic return

Study	Empirical Model: Offering causal capacity	Empirical Model: Agent cognitive frames	Empirical Model: Agent actions	Empirical Model: Collective outcome	Offered explanation & corresponding mechanism
					[M1a, M2a]
Guo et al. (2018)	n/a	Perceived effectiveness of: feedback mechanism, seller protection, cross-border delivery Perceived national integrity of buyers Seller's trust in buyers; perceived risk of chargeback fraud	Seller's intention to trade	n/a	Installment of institutional mechanisms (e.g., feedback, seller protection) → common understanding of how things are done & signal trustworthiness of buyer → control risks& produce trust [M1a, M2a]
Hinz et al. (2016)	TV (disaster, world cup, US presidential election)	n/a	n/a	Sales in unit	Other - External shock
Hong et al. (2018)	Social media activity; network embeddedness; public good vs. private good campaign	n/a	n/a	Contribution toward public goods	Embeddedness → trust & social image concerns → contribution toward public goods [M1a, M2a]
Hong and Pavlou (2017)	Service provider's country (country differences, IT development level, service provider's reputation (moderator)	n/a	Buyer selection (contract)	n/a	Providers' country → perceived technical competency and functional competency → buyer selection [M1a, M2a]
Hu et al. (2017)	Product price; mean; review volumes; SD of ratings; sales rank; price	n/a	n/a	Sales rank, rating	Review → awareness of self-selection bias & product quality perception → purchasing [M1a, M2a]
Huang et al. (2017)	Social capital (cognitive capital, structural capital, relational capital)	Satisfaction (economic satisfaction, social satisfaction); Perceived effectiveness of e-commerce institutional	Loyalty to the platform (purchase intention)	n/a	Social interaction between buyers and sellers → social capital → buyer satisfaction → loyalty [M1a, M2a]

Study	Empirical Model: Offering causal capacity	Empirical Model: Agent cognitive frames	Empirical Model: Agent actions	Empirical Model: Collective outcome	Offered explanation & corresponding mechanism
		mechanism (moderator)			
Jiang et al. (2018)	Displayed lending statistics; Non-salient lending information; Platform attributes as moderator (platform awareness, market share, participants composition)	n/a	n/a	Herding (cumulative investors by the end of a given week)	n/a
Kim and Ahn (2007)	Market maker reputation; market maker web usability & security; seller expertise & reputation	Trust in market maker; trust in sellers	Transaction intention	n/a	Company reputation/ web usability/ web security/ seller reputation → trust → transaction intention [M1a, M2a]
Kuan et al. (2014)	Previous "buy" information; "like" information	Attitude toward the deal;	Intention to purchase	n/a	Social influence (normative & informational) → conformity, desire to gain social rewards → payment [M1a, M2a]
Li and Hitt (2008)	Long-term average review; short-term deviation of the review (review bias)	n/a	n/a	Sales	n/a
Li and Wu (2018)	Past sales; past social media WOM; experience vs. search goods	n/a	n/a	Subsequent incremental sales	Past sales as quality signal → observational learning (belief update) → payment social media WOM → awareness effect & reduce quality uncertainty → payment [M1a, M2a]
Lin et al. (2017)	Product network attributes (diversity & stability; incoming vs. outcoming; co-view vs. co-purchase); product characteristics (list price, review volume, review rating, past sales, inventory, bookmarks)	n/a	n/a	Sales	Network diversity in the outgoing network → distract attention incoming network → increased exposure network stability→ product search effort perception & perception about the associated product [M1a]

Study	Empirical Model: Offering causal capacity	Empirical Model: Agent cognitive frames	Empirical Model: Agent actions	Empirical Model: Collective outcome	Offered explanation & corresponding mechanism
Liu et al. (2014)	Free version offering; free version ranking; free version rating; Paid version rating; Hedonic app	n/a	n/a	Paid version download	Free version → awareness, interest, create direct interaction → better affective and cognitive responses → payment Ratings → product visibility and quality → payment [M1a, M2a]
Oestreicher-Singer and Sundararajan (2012)	Network effect (PageRank)	n/a	n/a	Distribution of revenue (Gini)	n/a
Ou et al. (2014)	Use of IM; message box; feedback system; Interactivity; presence	Swift guanxi; trust	Repurchase	n/a	Use of IT → sense of interactivity & presence → create a condition for building mutual understanding (trust & guanxi) → reduce product uncertainty and transaction risk perceptions → repurchase [M1a, M2a]
Özpolat et al. (2013)	Seal; search engine referral; direct traffic; session duration	n/a	n/a	Purchase conversion	Third-party verification→ reduce info asymmetries & engender trust→ purchase [M1a, M2a]
Pavlou and Gefen (2004)	n/a	Perceived effectiveness of (feedback mechanism, escrow services, credit card guarantees), trust in intermediary trust in sellers; perceived risk	Transaction	n/a	Market-driven reputation & expectation → transaction [M2a]
Pavlou and Gefen (2005)	Buyer's past experience; sellers past performance;	Community psychological contract violation; individual psychological contract violation;	Transaction	n/a	Seller information → buyer's psychological contract violation, → transaction buyer's psychological contract violation→trust, perceived risk [M1a, M2a]

Study	Empirical Model: Offering causal capacity	Agent cognitive frames	Agent actions	Collective outcome	Offered explanation & corresponding mechanism
		perceived effectiveness of institutional structures; trust in community of sellers; perceived risk, price premium			
Reiner et al. (2014)	Buy-now feature	n/a	Number of bids per bidder;	Number of bidders; auctioneer profits per auction	Buy-now feature → prevent high loss → attract additional bidders, number of bids per bidder [M1a, M2a, M5]
Thies et al. (2016)	Past eWOM (Facebook shares, comments); popularity (other's funding behavior)	n/a	Current eWOM; Contribution decision	n/a	eWOM → pre-choice evaluation → contribution decision contribution→ self-representation, self-enhancement → diffuse eWOM other backers' funding behavior →infer product utility → funding decision [M1a, M2a]
Thies et al. (2018)	Input control	n/a	n/a	Fund demand, supply and concentration	n/a
Van Slyke et al. (2006)	N/A	Concern for information privacy; familiarity; Trust; risk perception	Willingness to transact	n/a	Merchant relationships as social contract →Privacy violation (concern) → risk & trust → transaction willingness [M2a]
Wu et al. (2013)	Review volume; review valence; product price	Buyer's risk attitude (risk-averse, risk-neutral, risk-seeking)	Willingness to pay (WTP)	n/a	Review volume→ buyer's perceived risk→ WTP review valence → buyer's perceived risk & perceived value→WTP price → buyer's perceived risk & perceived value→WTP [M1a, M2a]

Study	Empirical Model				Offered explanation & corresponding mechanism
	Offering causal capacity	Agent cognitive frames	Agent actions	Collective outcome	
Xu and Chau (2018)	Amount of direct lender-borrower communication; Positive vs. negative comments; Response timeliness; Credit grade (moderator)	Perceived accuracy; perceived completeness;	n/a	The listing's funding success; The final interest rate	Information source serves as quality signal → lender trust → lending decision Peer comments serve as quality signal →reduce risk and uncertainty → lending decision Information quality perception → risk assessment and trust → lending decision [M1a, M2a]
Zheng et al. (2018)	Entrepreneur activeness; sponsor co-creation; social connection	Perceived control; Self-rated investment; Intimate knowing; psychological ownership	commitment	n/a	Value cocreation (sponsored & autonomous) → psychological ownership → commitment [M1a, M2a]

References

Benlian, A., Titah, R., and Hess, T. 2012. "Differential Effects of Provider Recommendations and Consumer Reviews in E-Commerce Transactions: An Experimental Study," *Journal of Management Information Systems* (29:1), pp. 237-272.

Burtch, G., Ghose, A., and Wattal, S. 2013. "An Empirical Examination of the Antecedents and Consequences of Contribution Patterns in Crowd-Funded Markets," *Information Systems Research* (24:3), pp. 499-519.

Burtch, G., Ghose, A., and Wattal, S. 2016. "Secret Admirers: An Empirical Examination of Information Hiding and Contribution Dynamics in Online Crowdfunding," *Information Systems Research* (27:3), pp. 478-496.

Burtch, G., Hong, Y., and Liu, D. 2018. "The Role of Provision Points in Online Crowdfunding," *Journal of Management Information Systems* (35:1), pp. 117-144.

Carmi, E., Oestreicher-Singer, G., Stettner, U., and Sundararajan, A. 2017. "Is Oprah Contagious? The Depth of Diffusion of Demand Shocks in a Product Network," *Management Information Systems Quarterly* (41:1), pp. 207-221.

Chen, H., De, P., and Hu, Y. J. 2015. "It-Enabled Broadcasting in Social Media: An Empirical Study of Artists' Activities and Music Sales," *Information Systems Research* (26:3), pp. 513-531.

Feller, J., Gleasure, R., and Treacy, S. 2017. "Information Sharing and User Behavior in Internet-Enabled Peer-to-Peer Lending Systems: An Empirical Study," *Journal of Information Technology* (32:2), pp. 127-146.

Forman, C., Ghose, A., and Wiesenfeld, B. 2008. "Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research* (19:3), pp. 291-313.

Ge, R., Feng, J., Gu, B., and Zhang, P. 2017. "Predicting and Deterring Default with Social Media Information in Peer-to-Peer Lending," *Journal of Management Information Systems* (34:2), pp. 401-424.

Ghose, A., Smith, M. D., and Telang, R. 2006. "Internet Exchanges for Used Books: An Empirical Analysis of Product Cannibalization and Welfare Impact," *Information Systems Research* (17:1), pp. 3-19.

Gleasure, R., O'Reilly, P., and Cahalane, M. 2017. "Inclusive Technologies, Selective Traditions: A Socio-Material Case Study of Crowdfunded Book Publishing," *Journal of Information Technology* (32:4), pp. 326-343.

Gregg, D. G., and Walczak, S. 2008. "Dressing Your Online Auction Business for Success: An Experiment Comparing Two Ebay Businesses," *MIS Quarterly* (32:3), pp. 653-670.

Gu, B., Park, J., and Konana, P. 2012. "Research Note-the Impact of External Word-of-Mouth Sources on Retailer Sales of High-Involvement Products," *Information Systems Research* (23:1), pp. 182-196.

Guo, S., Guo, X., Fang, Y., and Vogel, D. 2017. "How Doctors Gain Social and Economic Returns in Online Health-Care Communities: A Professional Capital Perspective," *Journal of Management Information Systems* (34:2), pp. 487-519.

Guo, Y., Bao, Y., Stuart, B. J., and Le-Nguyen, K. 2018. "To Sell or Not to Sell: Exploring Sellers' Trust and Risk of Chargeback Fraud in Cross-Border Electronic Commerce," *Information Systems Journal* (28:2), pp. 359-383.

Hinz, O., Hill, S., Kim, J.-Y., Darmstadt, T. U., Karlsruhe Institute of, T., and Microsoft, R. 2016. "Tv's Dirty Little Secret: The Negative Effect of Popular Tv on Online Auction Sales," *MIS Quarterly* (40:3), pp. 623-644.

Hong, Y., Hu, Y., and Burtch, G. 2018. "Embeddedness, Prosociality, and Social Influence: Evidence from Online Crowdfunding," *MIS Quarterly* (42:4), pp. 1211-1224.

Hong, Y., and Pavlou, P. A. 2017. "On Buyer Selection of Service Providers in Online Outsourcing Platforms for It Services," *Information Systems Research* (28:3), pp. 547-562.

Hu, N., Pavlou, P. A., Zhang, J., University of Texas, A., Stevens Institute of, T., and Temple, U. 2017. "On Self-Selection Biases in Online Product Reviews," *MIS Quarterly* (41:2), pp. 449-471.

Huang, Q., Chen, X. Y., Ou, C., Davison, R. M., and Hua, Z. S. 2017. "Understanding Buyers' Loyalty to a C2c Platform: The Roles of Social Capital, Satisfaction and Perceived Effectiveness of E-Commerce Institutional Mechanisms," *Information Systems Journal* (27:1), pp. 91-119.

Jiang, Y., Ho, Y.-C., Yan, X., and Tan, Y. 2018. "Investor Platform Choice: Herding, Platform Attributes, and Regulations," *Journal of Management Information Systems* (35:1), pp. 86-116.

Kim, M.-S., and Ahn, J.-H. 2007. "Management of Trust in the E-Marketplace: The Role of the Buyer's Experience in Building Trust," *Journal of Information Technology* (22:2), pp. 119-132.

Kuan, K. K. Y., Zhong, Y., and Chau, P. Y. K. 2014. "Informational and Normative Social Influence in Group-Buying: Evidence from Self-Reported and Eeg Data," *Journal of Management Information Systems* (30:4), pp. 151-178.

Li, X., and Hitt, L. M. 2008. "Self-Selection and Information Role of Online Product Reviews," *Information Systems Research* (19:4), pp. 456-474.

Li, X., and Wu, L. 2018. "Herding and Social Media Word-of-Mouth: Evidence from Groupon," *MIS Quarterly* (42:4), p. 1331.

Lin, Z., Goh, K.-Y., Heng, C.-S., Nanjing, U., and National University of, S. 2017. "The Demand Effects of Product Recommendation Networks: An Empirical Analysis of Network Diversity and Stability," *MIS Quarterly* (41:2), pp. 397-426.

Liu, C. Z., Au, Y. A., and Choi, H. S. 2014. "Effects of Freemium Strategy in the Mobile App Market: An Empirical Study of Google Play," *Journal of Management Information Systems* (31:3), pp. 326-354.

Oestreicher-Singer, G., and Sundararajan, A. 2012. "Recommendation Networks and the Long Tail of Electronic Commerce," *MIS Quarterly* (36:1), pp. 65-83.

Ou, C. X., Pavlou, P. A., Davison, R. M., Tilburg, U., City University of Hong, K., and Temple, U. 2014. "Swift Guanxi in Online Marketplaces: The Role of Computer-Mediated Communication Technologies," *MIS Quarterly* (38:1), pp. 209-230.

Özpolat, K., Gao, G., Jank, W., and Viswanathan, S. 2013. "Research Note: The Value of Third-Party Assurance Seals in Online Retailing: An Empirical Investigation," *Information Systems Research* (24:4), pp. 1100-1111.

Pavlou, P. A., and Gefen, D. 2004. "Building Effective Online Marketplaces with Institution-Based Trust," *Information Systems Research* (15:1), pp. 37-59.

Pavlou, P. A., and Gefen, D. 2005. "Psychological Contract Violation in Online Marketplaces: Antecedents, Consequences, and Moderating Role," *Information Systems Research* (16:4), pp. 372-399.

Reiner, J., Natter, M., and Skiera, B. 2014. "The Impact of Buy-Now Features in Pay-Per-Bid Auctions," *Journal of Management Information Systems* (31:2), pp. 77-104.

Slyke, C., Shim, J. T., Johnson, R., Jiang, J., University of South, F., and University of Central, F. 2006. "Concern for Information Privacy and Online Consumer Purchasing," *Journal of the Association for Information Systems* (7:6), pp. 415-444.

Thies, F., Wessel, M., and Benlian, A. 2016. "Effects of Social Interaction Dynamics on Platforms," *Journal of Management Information Systems* (33:3), pp. 843-873.

Thies, F., Wessel, M., and Benlian, A. 2018. "Network Effects on Crowdfunding Platforms: Exploring the Implications of Relaxing Input Control," *Information Systems Journal* (28:6), pp. 1239-1262.

Wu, J., Arturo, E., and Gaytán, A. 2013. "The Role of Online Seller Reviews and Product Price on Buyers' Willingness-to-Pay: A Risk Perspective," *European Journal of Information Systems* (22:4), pp. 416-433.

Xu, J. J., and Chau, M. 2018. "Cheap Talk? The Impact of Lender-Borrower Communication on Peer-to-Peer Lending Outcomes," *Journal of Management Information Systems* (35:1), pp. 53-85.

Zheng, H., Xu, B., Zhang, M., and Wang, T. 2018. "Sponsor's Cocreation and Psychological Ownership in Reward-Based Crowdfunding," *Information Systems Journal* (28:6), pp. 1213-1238.

Appendix C – Correlation Matrix and Hyperparameters for Essay 3

Table C1. Correlation Matrix

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18
1-Title1	1.000																	
2-Title2	-0.648	1.000																
3-Ranking1	0.053	0.010	1.000															
4-Ranking2	-0.076	0.013	-0.625	1.000														
5-PhysicianTenure	0.350	-0.072	0.108	-0.146	1.000													
6-TotalDialogue	-0.038	0.051	0.026	-0.015	-0.032	1.000												
7-PatientPosts	-0.025	0.054	0.032	-0.020	-0.020	0.977	1.000											
8-PhysicianPosts	-0.064	0.027	0.001	0.004	-0.057	0.786	0.634	1.000										
9-SocialReturn	0.032	0.002	0.006	-0.020	0.081	0.254	0.233	0.242	1.000									
10-ServiceDuration	0.037	-0.017	0.010	-0.007	0.085	0.238	0.234	0.181	0.098	1.000								
11-ResponseRate	-0.063	-0.003	-0.040	0.035	-0.038	0.253	0.092	0.642	0.099	0.063	1.000							
12-QuestionFrequency	-0.007	0.006	0.016	-0.014	-0.029	-0.206	-0.148	-0.316	-0.180	-0.280	-0.311	1.000						
13-ServiceIntensity	0.113	-0.049	0.012	-0.010	0.095	0.116	0.120	0.071	-0.016	0.045	0.052	0.064	1.000					
14-OfflineConnection	-0.205	0.103	-0.004	0.035	-0.300	0.151	0.153	0.102	-0.133	-0.076	0.006	0.035	-0.080	1.000				
15-Status1	0.155	-0.077	0.003	-0.033	0.197	-0.115	-0.115	-0.081	0.075	-0.011	-0.028	0.039	0.036	-0.681	1.000			
16-Status2	0.047	-0.032	-0.005	0.006	0.103	-0.022	-0.036	0.026	-0.073	0.152	0.134	-0.123	0.104	-0.298	-0.139	1.000		
17-AnswerFrequency	-0.065	0.016	-0.020	0.013	-0.053	0.018	-0.040	0.179	-0.052	-0.195	0.498	0.376	0.051	0.041	0.008	-0.008	1.000	
18-Location	0.028	-0.018	0.033	-0.108	0.061	0.006	0.012	-0.013	0.018	0.009	-0.035	0.007	0.053	-0.096	0.074	0.049	-0.029	1.000

Note. Correlations above 0.3 are labeled in gray.

Table C2. Hyperparameter Selection and Tuning

ML Algorithm	Hyperparameters	Optimal Value	Explanation	Other Default Configurations
Naïve Bayes (H2O)	Laplace smoothing parameter [0-2]	1	It sets the conditional probability of a predictor.	- minimum standard deviation: 1e-10 - minimum probability for observations without enough data: 0.001
Simple decision tree	Minimum leaf size [2-10]	2	It defines the minimum sample needed in a leaf node, which is used to control overfitting.	- pruning: enabled (MDL method) - reduced error pruning: enabled - quality measure: Gini index - average split point (the split value for numeric features): the mean value of two adjacent feature values -tree depth: unlimited
	Minimum node size (min_leaf_size *2)	4	Minimum number of samples required to split a node. This is a stopping criterion; if the sample size is smaller this number, the tree will not grow further.	
MLP neural network*	Max iterations [100-600]	140	Maximum number of learning iterations. It describes the number of times a batch of data passes through the algorithm to complete one epoch.	- number of hidden layers: 1 - number of hidden neurons per layer: 10
Logistic regression	Prior distribution (uniform, Gauss or Laplace)	Gauss	The assumption of coefficient distribution which is related to regularization (i.e., the technique to shrink the learned regression coefficient towards zero to avoid overfitting). Uniform: no regularization Gauss: the coefficients are assumed to be normally distributed, which is equivalent to using L2 regularization. Laplace: the coefficients are assumed to follow a Laplace distribution, which is related to L1 regularization.	- algorithm: stochastic average gradient - learning rate: 0.1 - epsilon (determine whether the model converge): 1e-5
	Variance of prior distribution (*var*) [0.01, 0.05, 0.1]	0.1	Controls the degree of regularization. The regularization coefficient $\lambda = \frac{1}{var}$. Thus, the larger the variance, the less the regularization.	
	Epochs	10000	The maximum number of learning iterations. (The algorithm stops when it reaches convergence)	
Random forest (H2O)	Number of trees in the forest [50,300]	50	Including more trees is better to learn the data, but it significantly slows down the training process.	- minimum relative improvement rate: 0.00001 - sample rate per tree: 1 - class specific sample rate per tree: 1 - feature sampling rate per tree: 1 - feature sampling rate per split: 1
	Max number of levels in each decision tree (tree depth) [2,6]	6	The more splits a tree has, the more information can be captured from the data, but a higher depth means the algorithm learns the relations at a very specific level (i.e., overfitting).	
	Minimum number of observations in a leaf (min leaf size) [2-10]	2	It defines the minimum sample needed in a leaf, which is used to control overfitting.	

Table C2. Hyperparameter Selection and Tuning

ML Algorithm	Hyperparameters	Optimal Value	Explanation	Other Default Configurations
AdaBoost tree	Min leaf size [2-10]	8	The minimum sample needed in a leaf node.	- number of models to learn: 10 - Other default decision tree settings
Gradient Boost tree (H2O)	Number of trees [≤100, ≤200, ≤300]	100	Specifies the optimal number of trees to grow in the model.	- maximum tree depth: 10 - minimum observations for a leaf: 10 - loss function: auto - random seed: yes - - minimum relative improvement rate: 0.00001 - sample rate per tree: 1 - class specific sample rate per tree: 1 - feature sampling rate per tree: 1 - feature sampling rate per split: 1
	Learning rate [0.05, 0.1, 0.3, 0.5]	0.1	The scalar that determines the size of each learning step. A lower learning rate requires more trees to achieve the same level of fit.	
XGBoost tree (H2O)	Eta [0.01, 0.05, 0.1, 0.3, 0.5]	0.3	Similar to the concept of learning rate in gradient boosting.	- minimum sum of observation weight in a node: 1 - maximum delta step: 0 (no constraint) - subsampling rate: 1 - feature sampling rate per tree: 1 - feature sampling rate per split: 1 - L2 regularization (lambda): 1 - L1 regularization (alpha): 1 - tree construction method: auto - grow policy: split nodes closest to the root
	Gamma [0-10]	1	The minimum loss reduction required to make a further split of a tree	
	Tree depth [2-10]	6	The maximum depth of a tree.	
	Boosting rounds [50, 100, 200]	50	Rounds of boosting iteration.	

Note. (H2O) in brackets means the algorithms are run under H2O.ai frame, and the available hyperparameters are summarized at: http://docs.h2o.ai/h2o/latest-stable/h2o-docs/parameters.html. Other algorithms are run with KNIME platform.

Appendix D – Additional Analysis for Essay 3

1. Additional Analysis with Balanced Data

We randomly selected 492,902 free-only transactions (because N_{paid}=492,902) and repeated the whole procedures with three algorithms – logistic regression, decision tree and random forest. The model performance and feature importance results are presented in Table D1 and D2. In general, the changes for all performance measures are minor. Logistic Regression and Random Forest improved recall at about 4.5%, indicating a reduction in type II error, whereas the improvement in the decision tree is negligible. The change in overall performance measures (i.e., balanced accuracy, Cohen's kappa and AUC) is minor across all three algorithms. Feature importance ranking is very similar to the main analysis – offline connection, total dialogue, response rate, social return, prior exam and private are consistently ranked highly, whereas physician title, question frequency, and the second-tiered hospital ranking are consistently ranked low. However, the tree structure is much clearer for the balanced data (see Figure D1). In summary, features related to service quality and patient involvement seem to be more important than physician offline reputation (e.g., affiliation and title).

Table D1. Model Performance for Balanced Data

	LR		DT		RF	
	Score	Change	Score	Change	Score	Change
Recall	0.897	+0.046	0.955	+0.006	0.953	+0.045
Precision	0.905	+0.009	0.990	+0.002	0.966	-0.018
Specifity	0.906	-0.050	0.991	-0.004	0.968	-0.025
F-measure	0.901	+0.028	0.972	-0.004	0.959	+0.015
Accuracy	0.902	-0.021	0.973	-0.008	0.961	-0.006
Balanced accuracy	0.902	-0.001	0.973	+0.001	0.961	+0.010
Cohen's kappa	0.803	-0.015	0.946	-0.009	0.922	+0.001
AUC	1.000	0.000	0.988	0.000	0.989	+0.001
Optimal Hyperparameter	Prior distribution: Laplace Variance of prior distribution = 0.1 Epochs=10,000		Minimum leaf size= 4		Number of trees in the forest= 250 Minimum leaf size = 6 Tree depth = 6	

Table D2. Feature Importance Based on Balanced Data			
	LR	DT	RF
1	Response rate (-12.072)	Offline connection (1)	Offline connection (29.79%)
2	Offline connection (-5.025)	Social return (2)	Total Dialogue (19.14%)
3	Social return (-2.563)	Total dialogue (2)	PriorExam (16.20%)
4	Patient posts (-2.457)	Private (3)	Response rate (13.72%)
5	Total dialogue (2.340)	Response rate (3)	Patient posts (10.21%)
6	PriorExam (1.701)	PriorExam (4)	Social return (5.74%)
7	Private (-0.682)	Question_frq (5)	Answer_frq
8	Answer_frq (-0.283)	Answer_frq (5)	Private
9	Ranking 2 (-0.268)	Patient posts (6)	Question_frq
10	Question_frq (-0.137)	Title 1 (7)	Title 1
11	Title 1 (-0.099)	Ranking 2 (10)	Ranking 2

Note. LR-logistic regression. Coefficients are shown in the brackets. DT – (simple) decision tree. The numbers in the brackets show the highest level of tree splits; RF – random forest. The numbers in the brackets show the percentage feature importance (only those above 5% are shown).

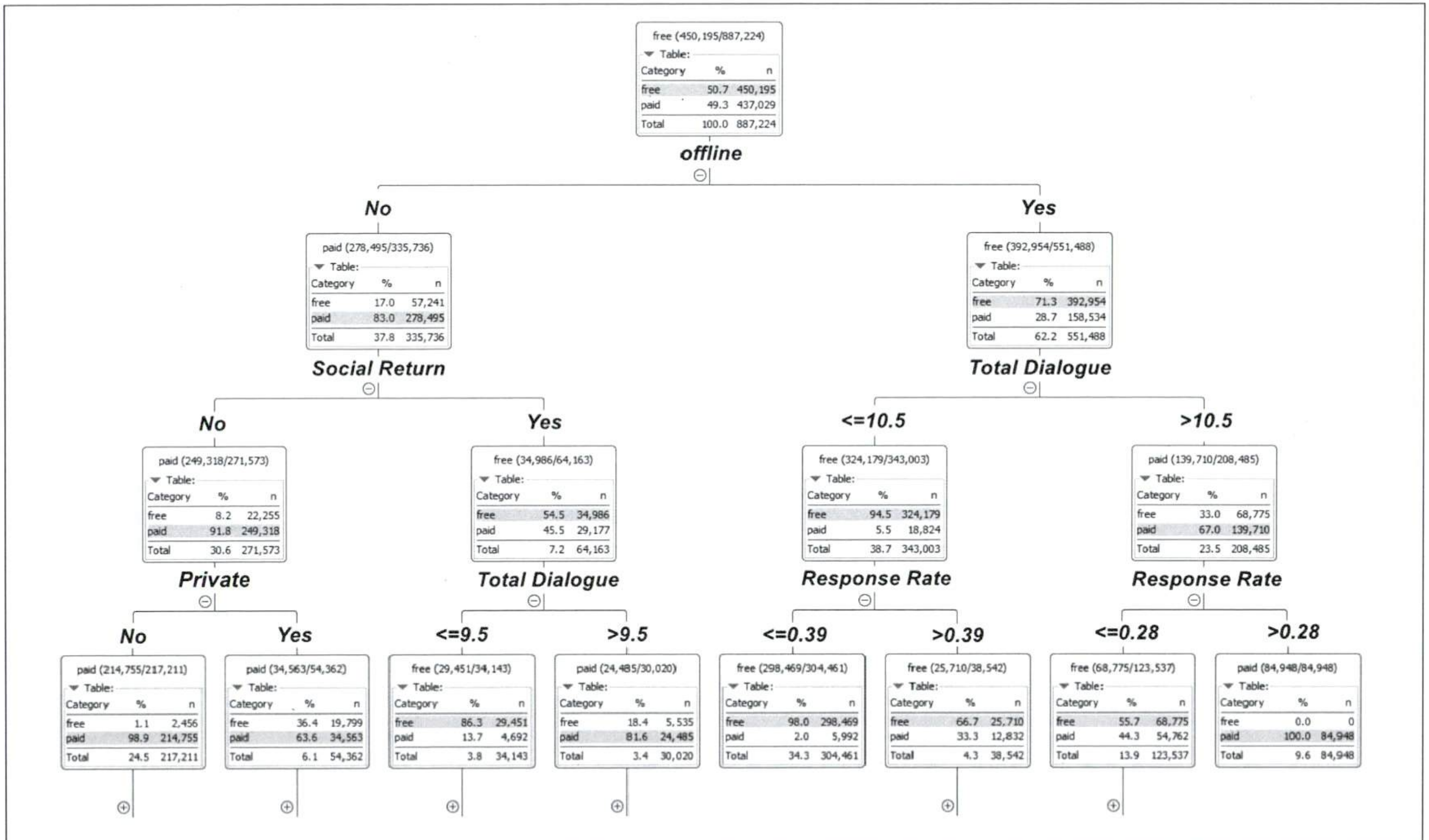


Figure D1. A Decision Tree Based on Balanced Data

2. Comparing Areas with Rich versus Few Healthcare Resources

We randomly selected 81,311 *free* records from remote areas and 474,203 *free* records from resource-rich areas to balance the *paid* records. Results are presented in Table D3 and D4. The model performance scores are similar to those of the main analysis (as well as the analysis with balanced data) with slightly lower scores for the remote areas, indicating that the applicability of our model is not influenced by physicians' location. However, although the overall feature importance ranking is similar to the main analysis (i.e., service quality features and patient involvement features rank much higher than physician reputation features), remote areas with few healthcare resources present quite different feature configurations according to the decision tree structure and logistic regression (see Figure D2). In general, the amount of total dialogue is the dominant feature that is associated with payment. For the consultations with fewer communication turns, patients with offline connections are less likely to pay, unless total dialogue and response rate are high. Social return may substitute payment as indicated by 73% of the patients who had no offline connections but provided social returns did not go beyond free services. However, the amount of patient posts only exhibits important impact for those from remote areas with few healthcare resources. It may be due to the differences in medical consultation habits, or difficulties in building trust and commitment for those "less known" providers. More research needs to be done to explore the reason behind this difference.

Table D3. Model Performance Comparison between Areas with Balanced Data												
	Areas Rich in Healthcare Resources						Areas with Few Healthcare Resources					
	LR (balanced)		DT (balanced)		RF (balanced)		LR (balanced)		DT (balanced)		RF (balanced)	
	Score	Change	Score	Change	Score	Change	Score	Change	Score	Change	Score	Change
Recall	0.898	0.002	0.956	0.001	0.954	0.001	0.928	0.031	0.948	-0.007	0.959	0.006
Precision	0.906	0.000	0.991	0.000	0.967	0.002	0.926	0.021	0.995	0.004	0.957	-0.009
Specifity	0.906	0.000	0.991	0.000	0.969	0.001	0.926	0.020	0.995	0.004	0.957	-0.011
F-measure	0.902	0.001	0.973	0.000	0.961	0.001	0.927	0.026	0.971	-0.002	0.958	-0.001
Accuracy	0.902	0.001	0.973	0.000	0.961	0.001	0.927	0.026	0.971	-0.002	0.958	-0.003
Balanced accuracy	0.902	0.001	0.973	0.000	0.961	0.001	0.927	0.026	0.971	-0.002	0.958	-0.003
Cohen's kappa	0.805	0.002	0.947	0.001	0.923	0.001	0.854	0.051	0.943	-0.003	0.916	-0.006
AUC	1.000	0.000	0.988	0.000	0.988	0.000	1.000	0.000	0.983	-0.005	0.989	0.000

Note. N_{rich}=1,482,554; N_{few}=100,010; $N_{few\text{-}paid}$ =81,311; $N_{rich\text{-}paid}$=474,203

Table D4. Feature Importance Comparison between Areas with Balanced Data						
	Areas Rich in Healthcare Resources			Areas with Few Healthcare Resources		
	LR	DT	RF	LR	DT	RF
1	Response rate (-12.02***)	Offline connection (1)	Offline (30%)	Response rate (-11.09***)	Total dialogue (1)	Total Dialogue (32.5%)
2	Offline connection (-5.11***)	Social return (2)	Total Dialogue (18.6%)	Offline connection (-4.12***)	Offline connection (2)	Offline (17.4%)
3	Social return (-2.69***)	Total dialogue (2)	Status1 (16.5%)	PriorExam (3.03***)	Response rate (2)	Response rate (16.8%)
4	Patient posts (-2.45***)	Private (3)	Response rate (13.6%)	Patient posts (-2.74***)	Social return (3)	Patient posts (16.60%)

Table D4. Feature Importance Comparison between Areas with Balanced Data						
	Areas Rich in Healthcare Resources			Areas with Few Healthcare Resources		
	LR	DT	RF	LR	DT	RF
5	Total Dialogue (2.33***)	Response rate (3)	Patient posts (9.9%)	Total Dialogue (2.60***)	PriorExam (4)	PriorExam (9.7%)
6	Status1 (1.62***)	PriorExam (4)	Social return (6%)	Status2 (1.32***)	Patient posts (5)	Question_frq
7	Status2 (-0.81***)	Answer_frq (6)	Answer_frq	Social return (-0.92***)	Title 1 (6)	Answer_frq
8	Ranking2 (-0.35***)	Question_frq (7)	Status2	Title1 (-0.49***)	Private (7)	Social return
9	Answer_frq (-0.31***)	Title1 (7)	Question_frq	Ranking2 (-0.18)	Question_frq (8)	Private
10	Question_frq (-0.13***)	Ranking2 (8)	Title1	Answer_frq (-0.06)	N/A	Title1
11	Title1 (-0.07***)	Patient posts (9)	Ranking2	Question_frq (-0.03)	N/A	Ranking2

Note. LR-logistic regression. Coefficients are shown in the brackets. DT – (simple) decision tree. The numbers in the brackets show the highest level of tree splits; RF – random forest. The numbers in the brackets show the percentage feature importance (only those above 5% are shown).

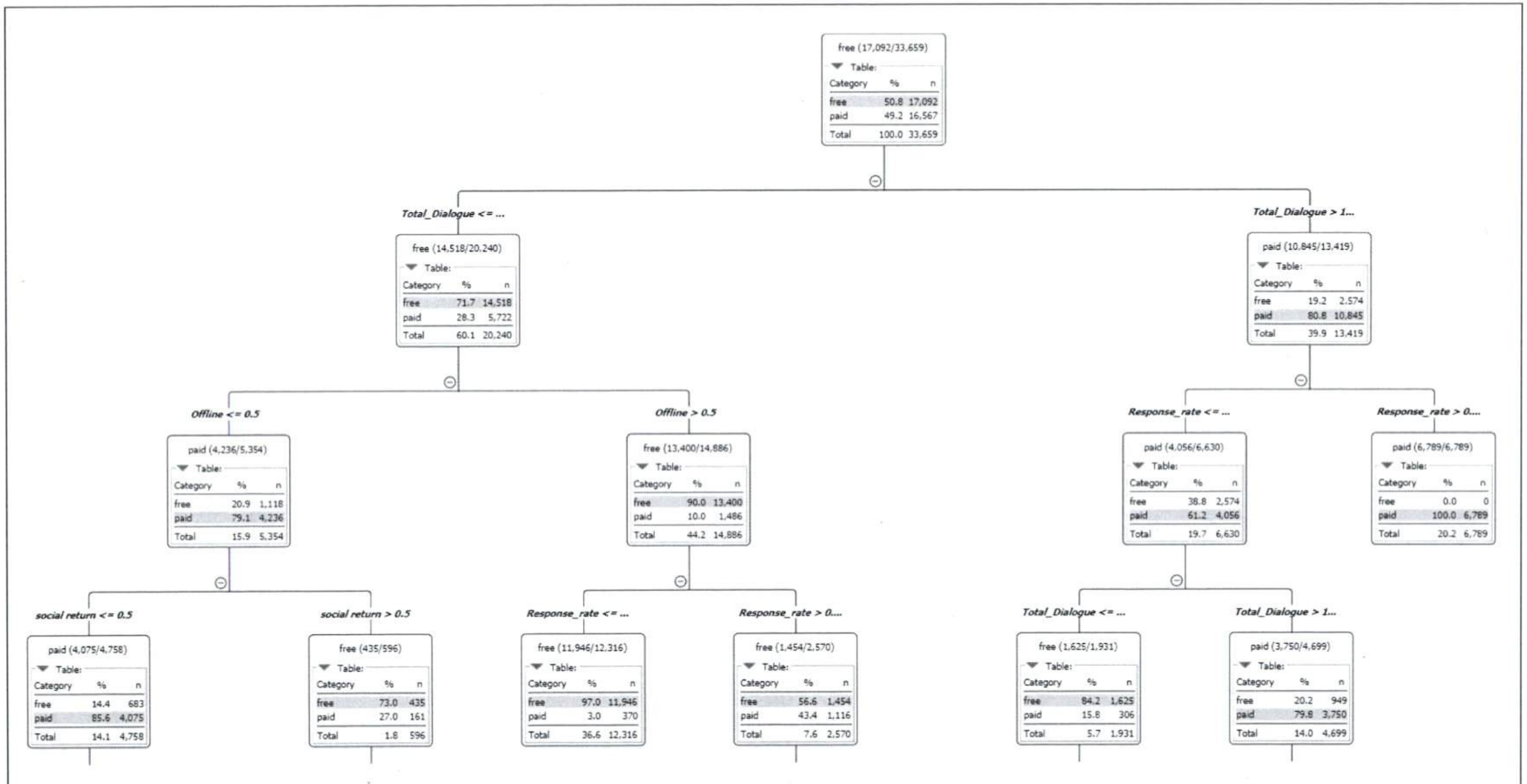


Figure D2. A Decision Tree with Balanced Data for Remote Areas with Few Healthcare Resources

3. Comparing the Ten-Year Model with a Four-Year Model

In total, there are 1,090,041 consultation records after 2015 (N_{free}=783,769 and N_{paid}=306,272). We create a balanced dataset by randomly selecting 306,272 free-only records. The ML performance results and feature ranking are presented in Table D5 and D6. Logistic regression exhibits a better classification for the more recent data – there is an 8.5% improvement in recall, 7.1% improvement in F-measure, 7% improvement in Cohen's kappa and 4.1% improvement in balanced accuracy – whereas decision tree has a minor reduction in performance. All three ML algorithms exhibit excellent overall classification performance. The feature ranking is consistent with the main analysis (as well as the other additional analyses) – total dialogue, offline connection, prior examination and response rate are ranked highly, and physician title and affiliation are ranked low. The amount of patient posts is a bit controversial since decision tree ranked it low, whereas the other two algorithms ranked it highly. In general, the performance of our model is not influenced by potential systematic differences due to the market cycle.

Table D5. Model Performance for Balanced Four-Year Data

	LR		DT		RF	
	Score	Change	Score	Change	Score	Change
Recall	0.936	0.085	0.954	0.004	0.954	0.046
Precision	0.951	0.055	0.989	0.000	0.974	-0.010
Specifity	0.952	-0.004	0.990	-0.006	0.975	-0.018
F-measure	0.944	0.071	0.971	0.002	0.964	0.019
Accuracy	0.944	0.021	0.972	-0.009	0.964	-0.002
Cohen's kappa	0.888	0.070	0.943	-0.012	0.929	0.008
Balanced accuracy	0.944	0.041	0.972	-0.001	0.964	0.014
AUC	1.000	0.000	0.986	-0.002	0.988	0.000

Table D6. Feature Importance Based on Balanced Four-Year Data

	LR	DT	RF
1	Total Dialogue (61.25***)	Offline connection (1)	Offline connection (30%)
2	Patient post (-45.27***)	Social return (2)	Total Dialogue (20%)
3	Response rate (-10.21***)	Total dialogue (2)	Response rate (17.2%)
4	Offline connection (-5.71***)	Private (3)	PriorExam (17%)

Table D6. Feature Importance Based on Balanced Four-Year Data			
	LR	DT	RF
5	PriorExam (2.09***)	Response rate (3)	Patient post (8.8%)
6	Social return (-1.43***)	PriorExam (4)	Private
7	Answer_frq (-0.71***)	Question_frq (5)	Social return
8	Private (0.36***)	Ranking 2 (7)	Answer_frq
9	Question_frq (0.15***)	Title 1 (8)	Question_frq
10	Ranking2 (-0.05)	Patient posts (9)	Title1
11	Title 1 (0.006)	Answer_frq (9)	Ranking2

Note. LR-logistic regression. Coefficients are shown in the brackets. DT – (simple) decision tree. The numbers in the brackets show the highest level of tree splits; RF – random forest. The numbers in the brackets show the percentage feature importance (only those above 5% are shown).

4. Additional Analysis with Outliers

There are 1,691,491 consultation records available in total if outliers are included (N_{paid}=545,134; N_{free}=1,146,357). A balanced dataset is created using 545,134 randomly selected free-only records. The performance measures and feature importance rankings presented in Table D7 and D8 exhibit similar patterns with only minor changes from the main analysis. This indicates that the model with 11 features is robust to outliers (e.g., extreme cases and possible bad data points due to deficiencies in the web crawler).

Table D7. Model Performance for Balanced Data with Outliers						
	LR		DT		RF	
	Score	Change	Score	Change	Score	Change
Recall	0.905	0.054	0.965	0.016	0.931	0.023
Precision	0.906	0.010	0.993	0.004	0.969	-0.015
Specifity	0.905	0.054	0.965	0.016	0.931	0.023
F-measure	0.905	0.032	0.979	0.010	0.950	0.005
Accuracy	0.905	-0.018	0.979	-0.002	0.951	-0.015
Cohen's kappa	0.811	-0.007	0.958	0.003	0.902	-0.018
Balanced accuracy	0.905	0.002	0.979	0.007	0.951	0.000
AUC	1.000	0.000	0.991	0.003	0.987	-0.001

Table D8. Feature Importance for Balanced Data with Outliers			
	LR	DT	RF
1	Total Dialogue (60.24***)	Offline connection (1)	Offline connection (24%)
2	Patient posts (-49.59***)	Social return (2)	Response rate (21%)
3	Response rate (-4.90***)	Total dialogue (2)	Total Dialogue (19%)
4	Offline connection (-3.91***)	Private (3)	PriorExam (14%)
5	Social return (-1.80***)	Response rate (3)	Patient posts (11%)
6	PriorExam (1.40***)	PriorExam (4)	Social return (6%)
7	Answer_frq (-0.67***)	Question_frq (5)	Question_frq
8	Private (-0.61***)	Answer_frq (8)	Private
9	Question_frq (-0.54***)	Title 1 (8)	Answer_frq
10	Ranking2 (-0.23***)	Patient posts (9)	Title1
11	Title 1 (-0.11***)	Ranking 2 (9)	Ranking2

Note. LR-logistic regression. Coefficients are shown in the brackets. DT – (simple) decision tree. The numbers in the brackets show the highest level of tree splits; RF – random forest. The numbers in the brackets show the percentage feature importance (only those above 5% are shown).