A Framework for Enhancing the Replicability of Behavioral MIS Research Using Prediction Oriented Techniques

Pratyush Nidhi Sharma
Associate Professor of Information Systems
Department of Information Systems, Statistics and Management
Culverhouse College of Business
The University of Alabama, USA
Email: pnsharma@ua.edu

Marko Sarstedt [a, b]
Professor of Marketing
[a] Ludwig-Maximilians-University Munich, Germany,
Email: sarstedt@lmu.de
[b] Babeș-Bolyai University, Romania

Christian M. Ringle [a, b]
Professor of Management and Decision Sciences
[a] Hamburg University of Technology (TUHH), Germany
Email: c.ringle@tuhh.de
[b] James Cook University, Australia,

Jun-Hwa Cheah
Associate Professor
Norwich Business School
University of East Anglia
Email: jackycheahjh@gmail.com/ J.Cheah@uea.ac.uk

Anne Herfurth
Instructor of Information Systems
Department of Information Systems, Statistics and Management
Culverhouse College of Business
The University of Alabama, USA
Email: aherfurth@ua.edu

Joseph F. Hair
Professor of Marketing
University of South Alabama, USA
Email: jhair@southalabama.edu

Acknowledgement

Even though this research does not use the statistical software SmartPLS (https://www.smartpls.com), Christian M. Ringle acknowledges a financial interest in SmartPLS.

A Framework for Enhancing the Replicability of Behavioral MIS Research Using Prediction Oriented Techniques

Abstract. The ongoing scientific discourse surrounding the replication crisis in behavioral research, including management information systems (MIS) research, underscores the importance of innovative and rigorous approaches to theory development and validation. This article proposes the 'EP-mixed' framework, which addresses the necessity of an ontological distinction between explanation and prediction in MIS theories, along with the epistemological challenges associated with conflating exploratory and confirmatory research during the design of robust, replicable theories. EP-mixed refers to theories that explain and predict (i.e., EP theories) developed using a mixed mode that combines the strengths of both exploratory and confirmatory research. The EP-mixed framework guides researchers in selecting appropriate analytical approaches based on their research goals and the type of theory being developed. While it can be applied in conjunction with a broad spectrum of statistical methods to enhance the robustness and replicability of MIS theories, we elaborate on the predictive analytic tools available in partial least squares structural equation modeling (PLS-SEM) as an exemplar for operationalizing the framework.

Keywords: Replicability; Exploration; Confirmation; Explanation; Prediction; EP-mixed; Open Science; PLS-SEM

1 Introduction

This article seeks to address the ontological and epistemological issues encountered by management information systems (MIS) researchers during novel theory development, particularly in light of recent concerns stemming from the non-replicability of many research findings. This concern over replicability, described as the *replication crisis*, was highlighted by an open science collaboration study which investigated the top three psychology journals and revealed that only 39% of their published findings were replicable (Open Science Collaboration, 2015). Critical inspections of factors associated with the replication crisis have become a focal point in broader social science research (Munafò et al., 2017; Nosek et al., 2018), management research (Hensel, 2021; Jensen et al., 2023; Pagell, 2021; Sarstedt et al., 2024), as well as in the MIS literature (Bogert et al., 2021; Dennis et al., 2020; Hu et al., 2023). Given the importance of these discussions, there is an acute need for a practical and robust framework that MIS researchers can utilize to enhance the replicability of their studies. We aim to fill this crucial gap.

Our approach is rooted in Gregor's (2006) taxonomy, which distinguishes between two fundamental ontological components of behavioral MIS theories: explanation and prediction. We further consider two significant epistemological factors argued to be contributing to the replication crisis in behavioral research: (1) the conflation of exploratory and confirmatory research (Munafò et al., 2017; Nosek et al., 2018), and (2) the lack of predictive validity and generalizability assessments in behavioral research (Hofman et al., 2017; Shmueli & Koppius, 2011; Ward et al., 2010; Yarkoni & Westfall, 2017). Building upon this discussion, we propose a framework that researchers can leverage to better position their work and choose appropriate modeling approaches. The proposed framework synthesizes and extends discourse scattered across disciplines to provide a practical and conceptual basis for enhancing replicability in MIS research.

We also highlight the crucial role of predictive analytic methods within this framework to facilitate the development of robust and replicable MIS theories. Predictive analytics has the potential to promote the development and validation of replicable socio-technical theories that generalize well (Hofman et al., 2017; Shmueli & Koppius, 2011; Ward et al., 2010; Yarkoni & Westfall, 2017). It is often seen as being synonymous with machine learning methods typically used in the computational social sciences, such as neural networks or random forests (Hofman et al., 2017; Hofman et al., 2021). However, behavioral MIS researchers can also harness the benefits of predictive analytics by employing structural equation modeling (SEM) methods such as partial least squares (PLS-SEM; Hair et al., 2022; Hair et al., 2024; Lohmöller, 1989; Wold, 1982), which follows a prediction perspective in the estimation of complex relationships between latent variables.

The primary purpose of this article is to propose a framework that researchers can utilize to conduct innovative, rigorous, and replicable research by incorporating predictive analytic techniques. The secondary purpose is to elucidate how predictive modeling techniques could be used in a PLS-SEM context for behavioral theory development and validation to enhance the replicability of MIS research. Our overall goal is to encourage the adoption of research practices that can enhance the replicability of MIS studies.

The article is structured as follows: First, we provide background on the nature and scope of theories and analytical perspectives in MIS and their role in the replication crisis. Next, we discuss the various theorization and analytical modes that researchers may employ and introduce a framework for enhancing replicability in behavioral MIS research. Subsequently, we showcase the tools available in PLS-SEM to operationalize the framework. Finally, we conclude with five key recommendations for authors and researchers leveraging our framework.

2 Background

MIS scholars are primarily dedicated to developing and testing theories that explain and predict information system use, adoption, and the associated individual and organizational-level outcomes in diverse contexts (Benbasat & Zmud, 2003; DeLone & McLean, 1992, 2003; Jeyaraj et al., 2023). They often assume a behavioral perspective and align themselves with the positivist paradigm to craft and validate quantitative models by drawing on data that is obtained from information systems users (Bariff & Ginzberg, 1982; Venkatesh et al., 2008). Their research usually starts with questions related to issues of substantive practical and theoretical relevance. The nature and scope of the theories cultivated for this purpose hinge on two fundamental considerations: (1) the overarching goal or *causa finalis*, which determines the ontological components of the theory under construction, and (2) the prevailing state of knowledge within the fast changing MIS field, which dictates the epistemological processes employed for theory construction and validation (Gregor, 2006).

The first consideration involves taking an ontological stance to theory building and asking *what* questions such as: "What are the ontological goals of the theory?" and "What are the main components of the theory?" In the past, the philosophy of science literature has discussed these issues at length, emphasizing that a theory could comprise of two core types of propositions: propositions for explanation and propositions for prediction (e.g., Dubin, 1969; Nagel, 1979; Popper, 1959). The management literature elucidated these aspects further (Sutton & Staw, 1995; Weick, 1989, 1995). MIS researchers have focused on the issue of developing native theories with theoretical relevance as well as practical relevance to practitioners (Benbasat & Zmud, 1999; Grover & Lyytinen, 2023; Lee et al., 2021; Straub, 2012; Straub & Ang, 2011), especially ones that address novel, emerging phenomena in the rapidly changing information technology landscape (Grover & Lyytinen, 2023).

To systematize the dispersed and limited discourse concerning the ontological aspects of MIS theories, Gregor (2006) introduced a structured taxonomy positing explanation and prediction as the two fundamental components of MIS theories. Depending on the theory's explanatory and predictive objectives, three types are delineated: theories designed solely for explanation (E), solely for prediction (P), and those encompassing both explanation and prediction (EP). While Gregor (2006) does not focus on empirically testing the components of explanation and prediction, subsequent MIS research has directly addressed this issue. For example, Shmueli (2010) distinguishes between explanatory modeling, which emphasizes in-sample evaluations (e.g., path significance, effect size, explained variance, model fit), and predictive modeling, which focuses on out-of-sample evaluations (e.g., predictive accuracy). The key difference in the approaches is that explanatory modeling focuses on how well a model fits and explains *observed* data to support ex post inferences, whereas predictive modeling focuses on how well a model generalizes to *unobserved* data to support ex ante inferences (Liengaard et al., 2021; Shmueli, 2010).

The second consideration relies on the epistemological perspective, which focuses on *how* theories are designed and validated, and poses questions such as: "Which modes of research are appropriate for constructing and testing theories?" or "Which methods and metrics should be used to judge the soundness of theories?" These issues have come under increasing scrutiny in the broader behavioral research domain due to concerns regarding the lack of replicability of seminal effects (Adler et al., 2023; Nosek et al., 2018). Scholars have identified two pivotal factors that significantly contribute to the ongoing replication crisis in research. The first factor pertains to the conflation of exploratory and confirmatory research, a matter extensively discussed by Nosek et al. (2018). The second factor concerns the inadequate predictive validity and generalizability of models employed in behavioral research, a deficiency underscored by prominent researchers such as Yarkoni and Westfall (2017), Shmueli and Koppius (2011), Hofman et al. (2017), as well as Ward et al. (2010). We discuss each below in turn.

First, the persistent conflation between exploratory and confirmatory approaches in research has been argued as the main cause of replication crisis (Munafò et al., 2017; Nosek et al., 2018). For example, researchers often explore many models in the background but only report those that increase their chances of publication (Petter & Hadavi, 2021). Such models are likely overfit to the data and may not be replicable. Thus, the practice of constructing, retrieving, and suppressing hypotheses after data has been collected and analyzed can be detrimental to the replicability of results (Hu et al., 2023). However, researchers can learn from the results of exploratory studies to legitimately inform confirmatory studies, and vice versa, if they do so transparently. This mixed approach requires a clear motivation, labeling, and separation of exploratory and confirmatory parts of the research and provides the potential for improved knowledge generation (Prosperi et al., 2019; Tukey, 1980). Open science advocates suggest a mixed mode approach that utilizes both a clearly labeled exploratory and a confirmatory approach (Nuzzo, 2014), although much like the larger sphere of management research, MIS has not adopted the mixed mode approach in the decade since it was first discussed. The lack of existing relevant guidelines detailing how a mixed mode approach should be applied in practice presents a challenge for researchers wishing to legitimately employ this approach.

Second, Yarkoni and Westfall (2017) and Ward et al. (2010) suggest that the replication crisis, really, is a crisis due to prior models' weaknesses in being able to predict unobserved data. They argue that instead of focusing on prediction, behavioral models are too often designed to fit observed data as accurately as possible using explanatory modeling. Similarly, Hofman et al. (2017, p. 486) note: "*Rather than asking whether a given theory can predict some outcome of interest, the accepted practice in social science instead asks whether a particular coefficient in an idealized model is statistically significant, and in the direction predicted by the theory.*"

In doing so, these models rely solely on in-sample metrics (e.g., model fit or variance explained) that simultaneously capture the replicable signal in the data as well as the nonreplicable noise idiosyncratic to the data (Pitt & Myung, 2002; Sharma et al., 2019). However, a good fit can mislead researchers into believing the model is adequately capturing the processes of interest and will necessarily predict well in other samples (Pitt & Myung, 2002). In fact, the best fitting models—in an explanatory sense—may not be the best predictive models (Sarstedt & Danks, 2022; Shmueli, 2010; Shmueli & Koppius, 2011). Hence, the sole reliance on explanatory modeling evaluated using in-sample metrics does not provide assurance that a model will be replicable in other samples (Ward et al., 2010). In essence, the prevailing argument here is that behavioral research needs to make itself more amenable to replication by making replicability a criterion for its models, such as by relying on predictive inference and model comparisons based on out-of-sample metrics and cross-validation to remove sample-specific noise (Killeen, 2019; Liengaard et al., 2021; Ward et al., 2010; Yarkoni & Westfall, 2017). However, existing guidelines on how predictive analytic techniques may be used in tandem with explanatory modeling are scarce.

This article is situated at the confluence of the dual issues identified as potential causes of the replication crisis and seeks to clarify the prevailing confusion regarding the objectives of exploratory versus confirmatory research. It also discusses the goals of explanatory versus predictive modeling, with a specific focus on the role of predictive modeling in theory confirmation and exploration using the composite-based PLS-SEM approach. Our emphasis on PLS-SEM as an exemplar stems from its robust explanatory and predictive capabilities (Cepeda Carrión et al., 2016; Richter et al., 2022) and its growing popularity in management disciplines for conducting both confirmatory and exploratory research (Benitez et al., 2020; Richter et al., 2016b; Sarstedt et al., 2022a). PLS-SEM also features prominently in the *International Journal of Information Management*, with 108 studies published between January 2020 and June 2023.

In order to do so, we utilize Gregor's (2006) taxonomy to propose the **EP-mixed Framework**,[1] an extended framework that considers the: (1) **theorization mode** (confirmatory, exploratory, or mixed), and the (2) **analytical mode** (explanatory, predictive, or a combination of both), that could be adopted by researchers in their studies. While the choice of the theorization mode is determined by the state of theoretical advancement and the novelty of phenomena being inspected (Rubin & Donkin, 2022), the choice of analytical mode is

[1] EP-mixed is in reference to theories that both explain and predict (i.e., EP-theories) developed using a mixed mode which combines the strengths of both exploratory and confirmatory research.

determined by the stated theoretical goals of the study (Shmueli, 2010). The straightforward EP-mixed framework elucidates the fundamental assumptions and prerequisites associated with each mode, enabling MIS researchers to position their studies accurately. This prevents any confusion between exploratory and confirmatory modes and ensures a solid foundation for selecting appropriate analytical modeling approaches. Additionally, the EP-mixed framework provides guidance on when and how researchers may employ predictive analytic techniques under various theory development scenarios. In doing so, it aims to tackle the crucial theory development challenge of establishing a clear distinction between exploratory and confirmatory modes, as well as addressing the deficiency of predictive validity assessments in MIS research. The EP-mixed framework enables MIS researchers to create and validate more robust, replicable, and practically relevant theories. The framework is method-agnostic, providing practical guidance on conducting a rigorous and replicable mixed mode study regardless of the statistical approach being used. We use PLS-SEM as an illustrative example, emphasizing that it is just one of several potential analytical methods that can be utilized to operationalize the EP-mixed framework. For example, Cho et al. (2019) have introduced predictive techniques in the context of generalized structured component analysis (GSCA; Hwang et al., 2023; Hwang & Takane, 2004, 2014)—also see Cho et al. (2023) and Cho et al. (2022b). Our extended framework may be used in conjunction with these techniques as well. The general principles of theorization and analytical modes underpinning the EP-mixed framework for replicable research are likely to be of interest to a substantial portion of behavioral MIS researchers.

3 Theorization mode: exploratory, confirmatory, or mixed?

The process of theorizing involves creating new theories or refining existing theories, as well as testing and validating them through empirical research and evidence. Theorization in behavioral management research is typically implemented using either the exploratory or a confirmatory mode (Richter et al., 2016b). We begin by examining the distinctions between exploratory and confirmatory modes and briefly discussing the problems that arise when they are conflated in practice. We note that, due to the replication crisis in behavioral research, this distinction has assumed greater significance for MIS researchers (Bogert et al., 2021; Hu et al., 2023). Subsequently, we discuss the potential of using these two modes in conjunction within a mixed mode (Nuzzo, 2014). By employing the mixed mode, researchers can harness the strengths of exploratory and confirmatory modes in tandem, ultimately facilitating the generation and validation of innovative and replicable MIS theories.

3.1 Exploratory versus confirmatory mode

The exploratory mode, characterized by its inductive nature and minimal dependence on existing theory, is employed for the imaginative and creative development of new ideas and concepts. This is often achieved by utilizing observed data with the aim of expanding the boundaries of existing knowledge. Phenomena detection is a critical aspect of the exploratory mode, which logically precedes theory development and validation (Jebb et al., 2017). Researchers in this mode seek to generate new hypotheses and propose innovative models about novel phenomena based on an ongoing interaction with observed data (Tukey, 1980). Consequently, exploratory research plays a vital role in pushing the boundaries of understanding and for addressing emerging questions and challenges (Adler et al., 2023; Richter et al., 2016a; Tukey, 1980).

In contrast, the confirmatory mode, characterized by its deductive nature and strong dependence on existing theory, entails the rigorous testing of theoretically motivated *a priori* hypotheses. The role of theory confirmation is to establish a robust foundation for scientific knowledge, serving as a vital bridge between theoretical constructs and observable phenomena. Popper (1959) clarified the role of confirmatory research in the ongoing development of scientific theories. He stressed the notion of falsification according to which theories can never be definitively confirmed but only falsified based on the accuracy of their predictions. While confirmatory research aims to validate theories, Popper (1959) suggests the true strength of a scientific theory lies in its ability to withstand repeated attempts at falsification. The role of theory confirmation, then, takes on a more nuanced character, serving not as a final confirmation but as a provisional corroboration contingent on the theory's resilience against falsification attempts. This approach emphasizes the continuous and dynamic nature of scientific inquiry, where theories are always subject to revision in the face of new evidence and challenges.

In the realm of scientific inquiry, theory exploration and confirmation are mutually reinforcing and indispensable, and contribute to a dynamic and progressive scientific landscape characterized by the interplay between innovation and validation (Jebb et al., 2017; Tukey, 1980). Weick (1989) viewed theory construction in management research as an art form of "disciplined imagination" that employs the concept of artificial selection, wherein researchers generate numerous models based on imagination, prior knowledge, and inductive logic. Subsequently, ideas surviving a particular test are retained for further confirmation. Weick's (1989) artificial selection process requires two essential components to be effective: (1) a process engendering diversity in ideas (i.e., exploratory research), and (2) a mechanism that distinguishes between less viable and more promising ideas (i.e., confirmatory research). In this paradigm, exploratory research serves to generate a diversity of ideas upon which confirmatory research functions as the discerning selection device. The iterative nature of this process is crucial, as each cycle contributes to the development of a more robust theory than the preceding one. If this strong mutual interdependency between exploratory and confirmatory research is utilized appropriately, it could result in major scientific advances (Tukey, 1980).

Conversely, "*when the two concepts [modes] get confused, and the two worlds collide, the fallout can be disastrous*" (Schwab & Held, 2020, p. 8). This collision of explanatory and confirmatory research is starkly highlighted in recent discussions surrounding the replication crisis in behavioral sciences (Shrout & Rodgers, 2018). Indeed, researchers' continued misclassification of their exploratory research as confirmatory research is a major cause of the replication crisis (Nosek et al., 2018). Using an achieved test result to rationalize a hypothesis and to subsequently argue that the result is a legitimate test of that hypothesis within the same dataset is a form of invalid circular reasoning (Rubin & Donkin, 2022).

The emergence of open science practices has intensified the focus on accurately distinguishing and appropriately categorizing confirmatory and exploratory research modes (Adler et al., 2023). Theoretical confirmation normally relies on deduction, with a researcher using established theory and logic to advance a priori hypotheses for testing. This method was designed with the expectation that the researchers would attempt to falsify a theory in good faith, which is an important facet of scientific conduct (Lakatos, 1976; Popper, 1959). Here, the critical point is that a purely confirmatory approach does not let the data speak (i.e., the role of data is strictly evidentiary) and is limited in its ability to generate novel insights that go beyond established theory and the researcher's logic (Tukey, 1980).

In contrast, the exploratory mode is inductive and relies on observed data to generate a new theory without a strong reliance on prior theories (i.e., the role of data is revelatory). Since researchers are not overly committed to a particular theory, the exploratory mode reduces researchers' self-fulfilling prophecy biases and facilitates inference to the best explanation

(Rubin & Donkin, 2022). However, because exploratory analyses are often unplanned and less reliant on theory, the social sciences have traditionally devalued exploratory research and venerated confirmatory research (Jacobucci, 2022; Jebb et al., 2017; Rubin & Donkin, 2022).

In practice, strictly confirmatory research tends to be restrictive and challenging, and researchers often resort to a mixed approach in which they manage an iterative dialog between existing theory and data to generate and test the new theory (Jöreskog & Wold, 1982; Richter et al., 2016a; Wold, 1982). In this process, researchers often explore multiple models in the background, but, due to publication constraints, they often mislabel their research as confirmatory and present their exploratory results selectively as if they were derived from a priori hypotheses (Adler et al., 2023; Petter & Hadavi, 2021). Even if researchers test and report different model configurations that adhere to their theoretical lens, they usually designate a single model as the winner, thereby ignoring the inherent uncertainty accompanying model comparisons (Preacher & Merkle, 2012; Rigdon et al., 2023; Sarstedt & Moisescu, 2023).

Open science practices strive to foster a clear delineation between confirmatory and exploratory facets, emphasizing transparency. Strictly speaking, research is classified as confirmatory when a researcher formulates a priori hypotheses and pre-analysis plans based on existing literature *before* undertaking the primary data collection (or, in the case of secondary or existing data, before analyzing the data), *and* proceeds to test these hypotheses without making any adjustments to the model or hypotheses once data collection has been completed. It is crucial to emphasize that a confirmatory approach is a one-shot approach that only gives researchers a single opportunity to evaluate the dataset (Gaus et al., 2015). Conversely, research is considered exploratory when the model is adapted on the basis of the results of the data analysis (Adler et al., 2023). The addition, removal, or modification of paths, constructs, variables, or indicators; or the utilization of modification indices to enhance the model's quality (e.g., model fit) through data analysis categorize a research study as exploratory, regardless of the statistical technique employed (Diamantopoulos, 1994; Sarstedt et al., 2014; Tomarken & Waller, 2003). For complex socio-technical and behavioral data, a mixed approach combining exploration and confirmation, while employing robust methodological safeguards, can yield insights more powerful than either approach alone (Prosperi et al., 2019; Tukey, 1980). We discuss below how a mixed approach can be utilized by researchers to build more robust and generalizable theories.

3.2 Exploration with confirmation: Towards a mixed mode of theorizing

Since theories do not appear out of thin air, both exploration and confirmation are required to generate and validate novel theories that will stand the test of time (Richter et al., 2016a). One method to encourage researchers to distinguish the confirmatory aspects of their research, while also permitting transparent reporting of any deviations from their initial research plan, is through preregistrations. Such exploratory deviations from the research plan are inevitable and have a tremendous potential to enhance the value of confirmatory research, if they are transparently communicated (Nosek et al., 2018).

In MIS research, the development of the Technology Acceptance Model (TAM) serves as a classic example illustrating where a two-phase mixed approach involving the exploratory and the confirmatory modes was employed to generate and validate a theory across multiple studies. TAM's initial formulation by Davis (1989) was based on exploring the concept of how users accept and adopt new information technologies. This initial phase involved the development of the theoretical framework and the generation of hypotheses about the factors influencing technology acceptance by means of surveys and interviews to gather insights into user perceptions and attitudes toward a variety of technologies' adoption. As TAM gained popularity, researchers began testing and validating it rigorously using large-scale surveys and quantitative analyses in a confirmatory mode. They used statistical techniques, such as SEM, to examine whether empirical data supported the proposed model and its relationships. Confirmatory research sought to validate the relationships between the TAM constructs by demonstrating the effects in a variety of contexts.

While this example demonstrates a mixed mode research program's success through theory development across a series of studies (Lakatos, 1976), there is no reason why a researcher shouldn't adopt a mixed mode for a single research study. Indeed, researchers advocate combining the exploratory and the confirmatory modes as part of a single study as long as they are clearly labeled (Nuzzo, 2015). For instance, when investigating novel phenomena, researchers might utilize the exploratory mode during the initial data analyses to flexibly generate insights and identify potential patterns or trends. The emphasis during this stage is on generating hypotheses without being distracted by the potential of false positives (Nuzzo, 2014). In the second stage, researchers might utilize the confirmatory mode while conducting rigorous hypothesis testing, using predefined methodologies and analysis plans to validate the initial findings from the exploratory stage. Key to this mixed mode approach is the use of a *separate sample* in this confirmatory stage. The confirmatory stage is meant to test the hypotheses generated in the first stage more rigorously and in a more controlled manner. This approach requires clear labeling of each stage and potentially preregistering the research project before its execution (Nosek et al., 2018). A mixed mode approach therefore enhances transparency and replicability, while retaining the study's innovative character, which are critical requirements for publication in top MIS journals.

The MIS literature has long argued in favor of integrating pluralistic approaches that cut across multiple research paradigms (e.g., Lee, 1991; Maier et al., 2023; Mingers, 2001, 2003; Robey, 1996). Combining different modes of research in a single study is already a well-established practice in MIS research. Mixed-method MIS research relies on a combination of hypothesis-generating approaches that are usually qualitative and hypothesis-testing approaches that are normally quantitative (Lee, 1991; Venkatesh et al., 2013; Venkatesh et al., 2016). For example, Califf et al. (2020) investigated the positive and negative effects of technostress on nurse appraisals and behaviors in a healthcare IT context. In the first phase, they conducted an exploratory qualitative study, using interviews with 32 nurses to generate hypotheses. In the second phase, they tested these hypotheses by using a survey of 402 nurses in a confirmatory study (i.e., using a separate sample from the first phase). A mixed mode approach would serve a similar purpose with respect to quantitative approaches, particularly when studying phenomena that are theory-weak, while maintaining *disciplined pluralism* to ensure that research approaches are chosen due to a study's goals and requirements and not because of researchers' subjective preferences (Robey, 1996).

Utilizing the strengths of exploratory and confirmatory research in a mixed mode approach can equip researchers to conduct powerful, innovative, and rigorous research. For this initial discussion, we are primarily concerned with single research articles that contain two or more studies, although these approaches can be used iteratively to generate a broader research program that encompasses many studies across a series of publications (e.g., Lakatos, 1976)—as we discuss later. Single research articles in such a paradigm represent interim attempts at theorizing that develop with time within a field (Weick, 1995).

Based on the current status of theoretical development in a field two distinct mixed mode approaches can be considered: (1) the *confirmatory-first* mixed mode and (2) the *exploratory-*

first mixed mode. The confirmatory-first mixed mode consists of a confirmatory study with well-defined research questions and robust support from existing literature as its first stage. This initial stage is followed by one or more post hoc exploratory studies. The primary objective of the confirmatory-first mixed mode approach is to test theoretically motivated hypotheses based on existing theories. This approach also aims to explore how these theories can be refined or extended through unplanned, intuitive, and data-driven tests, which may include exploring novel moderating or mediating effects, boundary conditions, or other factors. Post hoc results in a confirmatory-first mixed mode are considered tentative and provide material for confirmatory studies in the future. For example, drawing on past literature and findings, Hult et al. (2019) hypothesized about how the antecedents and consequences of customer satisfaction differ among customers when purchasing electronic goods online versus at a physical store in a confirmatory fashion. Subsequently, noting a lack of knowledge about the boundary conditions, they explored how these effects varied across customer demographics (gender, age, and education) and broader product categories. These post hoc exploratory results offer an initial indication of the extent of the theory's generalizability across different subgroups and lay a foundation for further confirmatory studies in the future.

In contrast, the first stage of the exploratory-first mixed mode approach comprises one or more exploratory studies. These studies commence with open-ended data analysis to guide the detection of novel phenomena, the generation of fresh hypotheses, or both. In the second stage of the exploratory-first mixed mode approach, the researcher leverages the knowledge generated from the exploratory stage to provide stronger, confirmatory evidence for the theory developed. This approach is valuable for situations where the research question is not yet well-defined or where the underlying processes are not well-understood. In MIS research, qualitative mixed-method researchers have successfully used the exploratory-first approach by utilizing an exploratory study followed by a confirmatory study. A typical goal in this research stream is phenomena detection and hypothesis generation using qualitative data in an exploratory fashion. Researchers can then validate the results through additional data collection in a confirmatory study (e.g., Califf et al., 2020). However, successful examples of exploratory-first mixed mode studies in quantitative behavioral research are scarce. Instead, this research stream has grappled with the issues of mislabeling exploratory results as confirmatory and confirming the results based on the same sample used to generate the hypotheses, raising concerns about *p*-hacking

(Sarstedt & Adler, 2023) and hypothesizing after the results are known, also known as HARKing (Hu et al., 2023). These questionable research practices present a substantial challenge for building a replicable body of research (Nosek et al., 2018).

The key to blending exploratory and confirmatory research effectively is to use them in a complementary manner, such as the proposed confirmatory-first and exploratory-first mixed mode approaches, with certain guardrails. Exploratory findings can be used to inform the development of hypotheses, and confirmatory findings can be used to refine theories. This iterative process can lead to a more comprehensive and nuanced understanding of the phenomena under investigation (Tukey, 1980). A successful application of an exploratory-first or confirmatory-first mixed mode approach requires the following important guardrails to be observed: (1) each stage must be clearly and distinctly labeled, (2) the analytical steps taken during exploratory analyses must be transparently and openly communicated, (3) the confirmatory study must follow an a priori analysis plan and one-shot approach to analyzing the data, and (4) the sample used to generate hypotheses must not be the same as the sample used to subsequently confirm the same hypotheses. Preregistrations can aid transparency in this process by helping researchers communicate their hypotheses, design choices, and analytical approaches before data are analyzed (Nosek et al., 2018). We adapt Schwab and Held's (2020) simple framework and expand it to offer additional guidelines specific to each theorization mode in Table 1 below.

	Confirmatory mode	**Exploratory mode**	**Confirmatory-first mixed mode**	**Exploratory-first mixed mode**
Overarching Goal	Confirm an existing theory	Generate a theory	*Stage 1:* Confirm an existing theory *Stage 2:* Enhance the theory	*Stage* 1: Generate a theory *Stage 2:* Confirm the generated theory
Reliance on Theory	Strong	Weak	Strong	Weak
Hypotheses	Clear, strong theory-based hypotheses Hypotheses need to be confirmed	No hypotheses or vague/uncertain hypotheses based on intuition or weak theory Hypotheses need to be generated	Clear, strong theory-based hypotheses Hypotheses need to be confirmed; further hypotheses need to be generated	No hypotheses or vague/uncertain hypotheses Hypotheses need to be generated and then confirmed
Purpose	Hypotheses confirmation	Hypotheses generation	*Stage 1:* Hypotheses confirmation *Stage 2:* Hypotheses generation	*Stage 1:* Hypotheses generation *Stage 2:* Hypotheses confirmation
Risks	Minimizing false positives	Minimizing false negatives	*Stage 1:* Minimizing false positives *Stage 2:* Minimizing false negatives	*Stage 1:* Minimizing false negatives *Stage 2:* Minimizing false positives
Suitability	Providing strong evidence to confirm theoretical expectations	Uncovering new insights to generate novel theory	Providing strong evidence to confirm theoretical expectations; subsequently making gradual enhancements to existing theory	Uncovering new insights to generate novel theory; subsequently providing strong evidence to confirm theoretical expectations
Data	Data is mute (i.e., its role is strictly evidentiary)	Data speaks (i.e., its role is strictly revelatory)	*Stage 1:* Data is mute *Stage 2:* Data speaks	*Stage 1:* Data speaks *Stage 2:* Data is mute
Analysis	One-shot and disciplined (will either confirm or disconfirm)	Iterative and intuitive	*Stage 1:* One-shot and disciplined *Stage 2:* Iterative and intuitive	*Stage 1:* Iterative and intuitive *Stage 2:* One-shot and disciplined
Sample	Single sample	Single sample	One or two samples or subsamples: *Stage 1:* Sample A for confirmation *Stage 2:* Sample B for exploration; or Sample A if not confirming	Two samples or subsamples: *Stage 1:* Sample A for exploration *Stage 2:* Sample B for confirmation
Evidence	Support in favor of hypotheses is considered *less* tentative. *Higher* confidence in implications of the study. Further confirmatory evidence may continue to be collected, especially in other contexts (theories can always be falsified and boundary conditions can be identified).	Support in favor of hypotheses is considered *more* tentative. *Lower* confidence in implications of the study. Further confirmatory evidence is required.	*Stage 1:* Less tentative support. More confidence in implications. *Stage 2:* More tentative support. Lower confidence in implications.	*Stage 1:* More tentative support. Lower confidence in implications. *Stage 2:* Less tentative support. Higher confidence in implications.

Table 1: Theorization modes

4 Analytical mode: Explanatory, predictive, or both?

The distinction between explanation and prediction has been widely acknowledged in the philosophy of science (e.g., Carnap, 2009; Douglas & Magnus, 2013; Douglas, 2009; Hempel, 1962; McCain, 2022), computational social sciences (Hofman et al., 2017; Hofman et al., 2021), and increasingly so in the behavioral social sciences (e.g., Liengaard et al., 2021; Shmueli & Koppius, 2011; Yarkoni & Westfall, 2017). In essence, this body of work shows that while explanation and prediction mutually facilitate one another, they are not empirically symmetrical (Douglas, 2009; Shmueli & Koppius, 2011). This analytical asymmetry manifests in the form of data and metrics suitable for testing explanatory and predictive hypotheses. However, despite the increasing acknowledgement of this distinction, the utilization of predictive analytic techniques remains infrequent in behavioral MIS research.[2] Researchers persist in making prospective inferences based on retrospective explanatory modeling (Hair & Sarstedt, 2021; Sarstedt & Danks, 2022). In the following sections, we discuss the differences in the analytical modes suited for testing explanation and prediction and how these two modes may be merged.

4.1 Explanatory versus predictive mode

As indicated above, Gregor's (2006) taxonomy delineates three primary types of MIS theories based on the ontological goals of explanation and prediction. First, theories for explanation (E) elucidate how and why certain phenomena occur. Our interpretation of E theories is broader. While Gregor (2006) exclusively considers qualitative approaches such as case studies, ethnographies, and hermeneutics as suitable for developing E theories devoid of testable propositions, we include behavioral theories offering explanations with testable propositions in the E theory classification as well. In the context of behavioral MIS research, which addresses socio-technical phenomena of a probabilistic nature (unlike natural phenomena driven by physical laws), these explanations are articulated as theoretical assertions. They posit that if specific preconditions are met (e.g., an individual perceives a software as easy to use), then certain outcomes (e.g., the individual's actual use of the software) are more likely to manifest in the *observed* data (Hempel, 1962). Further, according to Shmueli (2010), hypotheses tested using in-sample metrics are explanation-oriented rather than prediction-oriented. Thus, when utilizing SEM techniques, model fit metrics that test the discrepancy between model-implied correlations

[2] Our examination of PLS-SEM articles published in the *International Journal of Information Management* from January 2020 to June 2023 revealed that, while 30 studies (27.8%) out of a total of 108 studies provided the causal-predictive capability of PLS as a justification for its use in the study, only 8 studies (7.4%) had conducted out-of-sample prediction assessments (e.g., Bawack et al., 2023; Chen et al., 2023; Kumar et al., 2023). This finding presents further evidence of a limited adoption of predictive analytic techniques in MIS research, despite the accessibility of such tools in statistical software.

(or covariances) and observed correlations (or covariances) are geared toward explaining the *observed* data, but not unobserved data (Shmueli et al., 2016).

Second, theories for prediction (P) specify the *what* and *when* but not necessarily the *why* of an observed phenomenon (Carnap, 2009). Following Shmueli (2010), we define the term prediction not in its literal sense (i.e., predicting the future) but in a very specific and narrow sense: out-of-sample predictions of *unobserved* data, allowing for testing the generalizability or replicability of a model. Unobserved data here refers to data that a model has not seen and has not been accommodated to. This definition has two implications: First, it explicitly rules out testing observed data using in-sample metrics such as explained variance (R^2) for predictive purposes. Instead, prediction assessments rely on methods like cross-validation and metrics that focus on quantifying out-of-sample prediction errors (Shmueli & Koppius, 2011). This is an important distinction because MIS research has often erroneously employed in-sample metrics like R^2 to make predictive inferences. Second, based on the timing of when the observed and unobserved data were collected, predictive assessments can be conducted retrospectively (unobserved data collected in the past), cross-sectionally (observed and unobserved data collected simultaneously), or prospectively (unobserved data collected in the future). Retrospective predictions are useful for understanding past events and may account for temporal effects if the observed and unobserved data are historically separated in time. Cross-sectional predictions are useful for understanding current events and, by definition, do not account for temporal effects because all data (i.e., observed as well as unobserved) are collected at a single point in time. Finally, prospective predictions involve forecasting future events and explicitly separating the collection of observed exogenous and unobserved endogenous variables in time. The out-of-sample predictive accuracy of a model on the unobserved data signals its generalizability and replicability in a way that in-sample explanatory modeling cannot (Yarkoni & Westfall, 2017).

Third, theories for explanation and prediction (EP) encompass both the *how* and *why*, as well as the *what* and *when*. Gregor (2006) highlights that many MIS theories aspire to achieve both explanation and prediction. For instance, TAM (Davis, 1989) and DeLone and McLean's (1992) model of information system success are examples of theories that seek to both explain and predict. That is, the goal is to not only explain the relationship between independent and dependent variables in the observed data but also to predict unobserved data. For instance, a given hypothesis or proposition, such as “individuals’ perceptions of ease of use leads to actual use of the software,” can be interpreted in two distinct ways (Dubin, 1969). First, an explanatory interpretation seeks to establish a relationship between observed variables and the underlying mechanisms driving that relationship. It involves asserting how one or more variables, the independent variables, influence another variable, the dependent variable. For example, in the case of software use, one would argue that the perception of ease of use provides a plausible explanation for individuals’ actual use of software, while accounting for other factors (Shiau et al., 2024). Second, the predictive interpretation focuses on making inferences about unobserved data by leveraging patterns in the observed data. It does not necessarily aim to explain underlying mechanisms or causes. In this case, one would utilize the ease-of-use perceptions of an individual to predict their actual use of the software. Importantly, these inferences are made for unobserved data. These two interpretations of the same hypothesis determine the ontological nature of the theory being created (E, P, or EP), the epistemic process of validation being utilized (in-sample, out-of-sample, or both), and the nature of claims that can be supported (retrospective or prospective) – see also Shmueli (2010). However, MIS theories typically aim to both explain and predict (Gregor, 2006). This necessitates the simultaneous development of both the explanatory and predictive aspects of our models.

4.2 Utilizing explanatory and predictive analytic modes simultaneously

Critical to the development of EP theories are the dual issues of theoretical and practical relevance. Explanatory modeling is geared toward addressing theoretical relevance (retrospective inferences) through the use of in-sample metrics, while predictive modeling focuses on assessing practical relevance (prospective inferences) using out-of-sample metrics for generalizability (Shmueli & Koppius, 2011).[3] MIS research aims to achieve both theoretical and practical relevance (Benbasat & Zmud, 1999; Grover & Lyytinen, 2023; Lee et al., 2021; Straub, 2012; Straub & Ang, 2011). Hence, the simultaneous use of in-sample explanatory modeling and out-of-sample predictive modeling is crucial in assessing the effectiveness of model performance in

[3] We acknowledge that *practical relevance* is a broad term with numerous potential definitions and considerations (Nicolai & Seidl, 2010). Given that MIS is an applied discipline, we specifically adopt the interpretation by Dubin (1976, p. 32), who emphasizes that "*From the standpoint of sheer utility, a theoretical model is best judged by the accuracy of the predictions generated by it.*" This definition focuses on the extent to which a model’s prescriptions (predictions) effectively apply to the real-world, making it useful to practitioners (Gallien et al., 2016).

developing EP theories (Shmueli & Koppius, 2011). Additionally, predictive analytic techniques such as cross-validation can aid in discovering new theoretical insights.

A straightforward approach to conducting this hybrid analysis is to complement the in-sample explanatory assessments with out-of-sample predictive assessments at each step of the analysis. This evaluation necessitates simultaneously assessing both the substantive significance (i.e., magnitude) and the statistical significance of changes in explanatory and predictive capabilities to determine their meaningfulness (Kelley & Preacher, 2012; Sullivan & Feinn, 2012). Table 2 below illustrates four cases that may arise when utilizing explanatory and predictive modeling concurrently in a stepwise approach.

		Change in the in-sample explanatory power	
		Not Meaningful	Meaningful
Change in the out-of-sample predictive power	Not Meaningful	**Case 1: Change is not relevant** - Does not improve explanation - Does not improve prediction	**Case 2: Change is theoretically relevant but may cause overfitting to noise** - Improves explanation - Does not improve prediction
	Meaningful	**Case 3: Change is practically relevant but suggests hidden explanatory mechanisms** - Does not improve explanation - Improves prediction	**Case 4: Change is theoretically and practically relevant** - Improves explanation - Improves prediction

Table 2: Coevolution of theoretical and practical relevance

We first consider *Case 1* in Table 2. If a proposed modification, such as the addition of a variable, moderator, or mediator, does not meaningfully alter the in-sample explanatory performance or out-of-sample predictive accuracy of the model, it suggests that the modification lacks both theoretical and practical relevance. In this case, the adjustment does not sufficiently enhance explanation or prediction and suggests a theoretical dead-end.[4]

[4] Alternatively, from a robustness check standpoint, if an additional variable fails to enhance either explanation or prediction, it indicates the robustness of the established model against the introduced change.

In *Case 2*, a meaningful change in the in-sample explanatory performance without a correspondingly meaningful change in out-of-sample predictive accuracy indicates an enhancement in theoretical relevance but not in practical relevance. While the adjustment fails to improve prediction, it does enhance the model's ability to explain the observed (training) data. However, caution is warranted, as this improvement may stem from the model overfitting to sample-specific noise (i.e., illusory trends) rather than capturing generalizable signals in the data, potentially compromising practical relevance and generalizability (Yarkoni & Westfall, 2017).

In Case 3, a modification results in improved out-of-sample predictive accuracy without affecting the in-sample explanatory performance. This improvement indicates enhanced practical relevance of the model; however, the lack of impact on explanatory power suggests that it may be influenced by hidden explanatory mechanisms. For instance, an introduced variable may not be statistically significant but still enhances the model's predictive accuracy, capturing associations crucial for predicting outcomes beyond the training data or under different conditions (Shmueli & Koppius, 2011). Such situations are not uncommon in social science models. For example, Ward et al. (2010) and Hill and Jones (2014) empirically demonstrate that the statistical significance of a variable and its subsequent ability to improve in-sample model fit is not necessarily related to its out-of-sample predictive ability. Similarly, Sarstedt and Danks (2022) empirically demonstrate the inverse relationship between a model's explained variance and its out-of-sample predictive accuracy.

Finally, *Case 4* represents the most favorable scenario in which a meaningful change occurs in both in-sample explanatory power and out-of-sample predictive power. This indicates that the modification is both theoretically and practically relevant, contributing to an improved understanding of the underlying dynamics and enhanced predictive capabilities. The improvement in the predictive abilities of behavioral models, coupled with their explanatory capacities, can have a substantial social impact (Hofman et al., 2021; Kleinberg et al., 2015).

5 Bringing it all together: The EP-mixed framework for enhancing replicability in behavioral MIS research

To advance the development of replicable EP theories in MIS research, we propose the EP-mixed framework that combines the elements of the theorizing and analytical modes discussed in the previous sections (Table 3). It empowers researchers to transparently situate their research by critically evaluating the rationale behind their choice of theorization and analytical modes.

Theorization \ Analytical	Explanatory (E) mode	Predictive (P) mode	Explanatory-Predictive (EP) mode
Confirmatory mode	Confirming in-sample model fit [a priori].	Confirming out-of-sample model predictions [a priori].	Confirming in-sample model fit and out-of-sample model predictions [a priori] using a one-shot analysis. Preregistration of an a priori analysis plan is recommended.
Exploratory mode	Improving in-sample model fit [a posteriori].	Improving out-of-sample model predictions [a posteriori].	Improving in-sample model fit and out-of-sample model predictions [a posteriori]. Preregistration of an analysis plan is recommended as phenomena detection or hypothesis generation.
Confirmatory-first mixed mode	Confirming in-sample model fit [a priori]; subsequently improving in-sample model fit using a separate (or same) sample [post hoc].	Confirming out-of-sample model predictions [a priori]; subsequently improving out-of-sample model predictions using a separate (or same) sample [post hoc].	Confirming in-sample model fit and out-of-sample model predictions [a priori]; subsequently improving in-sample model fit and out-of-sample predictions using a separate (or the same) sample [post hoc]. Preregistration of an a priori analysis plan for the confirmatory stage is recommended. Analysis should be one-shot. Further post hoc exploratory analyses should be clearly labeled and described as a deviation from the preregistered plan.
Exploratory-first mixed mode	Improving in-sample model fit [a posteriori]; subsequently confirming in-sample model fit using a separate sample [a priori].	Improving out-of-sample model predictions [a posteriori]; subsequently confirming out-of-sample model predictions using a separate sample [a priori].	Improving in-sample model fit and out-of-sample model predictions [a posteriori]; subsequently confirming in-sample model fit and out-of-sample predictions using a separate sample [a priori]. For the exploratory stage, preregistration of an analysis plan is recommended as phenomena detection or hypothesis generation, along with a clear labeling and description of the exploratory study. For the confirmatory stage, preregistration of an a priori analysis plan is recommended. Analysis should be one-shot.

Table 3: The EP-mixed Framework for Replicable Research

While the EP-mixed framework spans a broad spectrum of theorization and analytical modes, we will provide an in-depth discussion of how the explanatory-predictive (EP) mode can be used in conjunction with the four theorization modes (shaded in gray in Table 3). We restrict our attention to the EP mode for two main reasons: (1) theories focusing solely on prediction are rare in behavioral MIS research and are more commonly found in the computational social sciences (Gregor, 2006), and (2) most, if not all, quantitative behavioral studies utilizing SEM techniques explicitly or implicitly aim to provide both explanation and prediction capabilities (Hair & Sarstedt, 2021; Shmueli & Koppius, 2011).

5.1 Confirmatory EP mode

The confirmatory EP mode requires the adoption of a meticulous approach as outlined below. Preregistration is highly recommended, emphasizing the importance of a transparent and predetermined analysis plan (Adler et al., 2023; Hu et al., 2023). Theory confirmation involves deductive logic, with researchers grounding their investigations in hypotheses derived a priori and adhering to a focused and pre-determined analytical trajectory (Munafò et al., 2017; Nosek et al., 2018). In a strict sense, the confirmatory EP mode involves a one-shot analysis based on a pre-analysis plan, with no alterations permitted based on the results. This stringent approach ensures the integrity of the confirmation process.[5]

The confirmatory EP mode itself encompasses both in-sample model fit and out-of-sample predictive accuracy assessments. In this dual evaluation, the model is scrutinized for its ability to surpass benchmarks or outperform alternatives, confirming its theoretical and practical relevance. To consider a hypothetical example, based on existing theory and deductive logic, a researcher proposes that behavioral intention mediates the effect of facilitating conditions on the actual use of a technology. A stepwise approach would be to conduct the established mediation tests and confirm improvement in model fit (Hair et al., 2022, Chapter 6). Subsequently, a predictive test could be conducted to assess the improvement in out-of-sample predictive accuracy of the proposed model. A meaningful improvement in the in-sample model fit implies the theoretical relevance of the proposed mediating effect, supporting retrospective inferences. Simultaneously, a meaningful improvement in the out-of-sample predictive power signifies practical relevance and generalizability, supporting prospective inferences (Shmueli & Koppius, 2011).

It is crucial to note that the analysis may yield one, both, or neither of these confirmations, emphasizing the nuanced nature of theory confirmation in empirical research (refer to Table 2 above). In the hypothetical example above, the analysis may confirm either the theoretical or practical relevance, or both, or neither, of the proposed mediator. While we illustrated one

[5] However, as we discuss below in our recommendations, deviations from a pre-analysis plan due to unanticipated factors can, and do, occur (Claesen et al., 2021). The extent of such deviation determines whether the study retains its confirmatory character (Nosek et al., 2018).

example of confirmation through mediation analysis, we acknowledge numerous possibilities for deductive modifications that researchers may employ for theory confirmation, including, but not limited to, additional variables, alternative paths or model structures, or moderators. In all cases, however, a successful application of the confirmatory EP mode demands strict adherence to the pre-analysis plan (or preregistration), a one-shot analysis, and transparent communication of the achieved results concerning both in-sample explanation and out-of-sample prediction.

5.2 Exploratory EP mode

The exploratory EP mode integrates in-sample explanatory modeling and out-of-sample predictive modeling to unravel phenomena, detect patterns, and generate theories. This mode is characterized by its intuitive and iterative nature, involving stepwise changes in an initial model, such as the addition of variables, paths, moderators, mediators, and indicators. Model comparisons play a pivotal role in emphasizing the coevolution of the model's explanatory and predictive power. The exploratory process adheres to inductive-abductive logic, fostering a dynamic dialog between theory and data (Tukey, 1980; Wold, 1982), with change actively encouraged based on results (Richter et al., 2016b). Here, out-of-sample tools such as cross-validation can play a dual role by acting as a guardrail against overfitting a model too tightly to the observed data, as well as documenting the improvements in the predictive accuracy of the model (e.g., Hofman et al., 2017; Shmueli & Koppius, 2011; Ward et al., 2010; Yarkoni & Westfall, 2017).

Unlike the confirmatory mode, a detailed preregistration is not possible in the exploratory EP mode since a strict analytical plan is not predetermined. However, researchers may still preregister their study and communicate their intention to perform exploratory research in advance (Fife & Rodgers, 2022; Szollosi et al., 2020). An explicit management of flexibility is necessary through transparent documentation of analytical steps taken, minimizing the risk of succumbing to the *garden of forking paths* and overfitting (Yarkoni & Westfall, 2017). Importantly, findings attained in the exploratory EP mode are considered provisional and should not be presented as confirmatory (Jebb et al., 2017).

While out-of-sample metrics offer additional confidence in exploratory findings than in-sample explanatory metrics alone, these findings are not validated until tested on a *separate* dataset using the confirmatory EP mode described above. The main distinguishing factor between the exploratory and confirmatory EP modes lies in the adherence to a pre-analysis plan based on a priori hypotheses (Nosek et al., 2018).

5.3 Confirmatory-first mixed EP mode

The confirmatory-first mixed mode embodies a two-stage research approach that amalgamates components from both confirmatory and exploratory EP modes, in that order, necessitating adherence to the distinct requirements of each. The initial stage of this mode entails a rigorous confirmatory study with a priori hypotheses and a pre-analysis plan that is ideally preregistered. In this phase, a researcher executes one-shot confirmatory tests, employing both in-sample explanatory modeling and out-of-sample predictive modeling. The outcomes of this test may affirm the theoretical relevance, practical relevance, both, or neither of the a priori hypotheses, as outlined in Table 2.

Following the confirmatory phase, the researcher initiates a post hoc exploratory study aimed at probing the boundary conditions of the proposed theory (e.g., via moderators), identifying mediating mechanisms, or enhancing either its explanatory or predictive power, or both. While the exploratory study may leverage the same dataset, any post hoc findings are considered tentative until corroborated in a subsequent study utilizing a separate sample. It is imperative that such post hoc tests are clearly labeled as exploratory and any deviations from the preregistered plan are explicitly noted (Nosek et al., 2018).

Although the confirmatory-first mixed mode is a commonly employed research approach, it is noteworthy that researchers have yet to utilize preregistrations to achieve its transparent implementation to enhance the replicability of findings (Adler et al., 2023). This indicates a gap in optimal utilization of the process despite its prevalence in research.

5.4 Exploratory-first mixed EP mode

The exploratory-first mixed mode represents a two-stage approach that integrates elements from both exploratory and confirmatory EP modes, in that order, and effectively meets the distinct requirements of each. Initiated with exploratory research, this mode aims to unveil new patterns, relationships, and mechanisms while generating potential hypotheses. This initial stage plays a pivotal role in stimulating creativity and refining both the in-sample model fit and out-of-sample model predictions a posteriori. Researchers meticulously adjust and refine the model based on examining data, incorporating predictive tools to guard against overfitting, and documenting improvements in predictive accuracy (Hofman et al., 2017; Shmueli & Koppius, 2011; Ward et al., 2010; Yarkoni & Westfall, 2017).

Having identified promising hypotheses, researchers progress to the confirmatory stage, guided by a preregistered plan whenever feasible. Drawing on insights gained during the exploratory phase, this stage rigorously tests hypothesized relationships (a priori), with the goal of confirming both in-sample model fit and out-of-sample model predictive accuracy using a *separate* sample. In this case, the segregation of samples between the two stages is paramount to ensuring the validity of the findings (Nosek et al., 2018). Predictive tools offer confirmation regarding whether a model surpasses prediction benchmarks or provides superior predictive accuracy compared to an alternative model (Liengaard et al., 2021). The second stage may affirm the theoretical relevance, practical relevance, both, or neither, as illustrated in Table 2.

The exploratory-first mixed mode not only agrees with Weick's (1989) view on theory development as *disciplined imagination* but also with the *strong inference* paradigm that aims to make rapid scientific progress (Platt, 1964). The integration of exploratory and confirmatory stages within the exploratory-first mixed mode enables researchers to capitalize on the strengths of both approaches, thereby fostering more robust EP theories. Additionally, preregistering the plan in advance for the confirmatory study enhances research transparency and replicability, instilling greater confidence in the findings (Nosek et al., 2018).

5.5 The EP-mixed framework cycle

The success of natural science disciplines is often attributed to a cyclical model where exploratory research with tentative findings leads to confirmatory research with more definitive claims that in turn open avenues for further exploratory research (Lieberman, 2020). In our initial conceptualization of the EP-mixed framework, our primary emphasis was on enhancing the replicability of individual studies. We now discuss the framework's potential contribution to the development of new, replicable research programs or the enhancement of existing ones in MIS research. This is achieved through a robust application of the confirmatory-first and exploratory-first mixed EP modes in an iterative fashion.

A research program, as articulated by Lakatos (1971), constitutes a series of studies conducted over time in a discipline to gain an in-depth understanding of a phenomenon or a set of related phenomena. An illustrative example in the MIS domain is the examination of antecedents of the adoption of new technologies in various contexts. Such programs are distinguished by a *core* theoretical foundation and a set of *auxiliary* hypotheses. The core encompasses facts and theoretical assertions generally accepted as true under most circumstances, though not universally, while auxiliary hypotheses remain open to modification and refutation, aiming to enrich the core (Lakatos, 1968). In the case of technology adoption literature, the theoretical linkages in TAM and Unified Theory of Acceptance and Use of Technology (UTAUT) serve as the core, with additional hypotheses seeking to extend or modify these models considered auxiliary (e.g., Blut et al., 2022). The core embodies knowledge solidified over time, while auxiliary hypotheses are adaptable and may be integrated into the core or replaced entirely.

In contrast to the naïve falsificationism proposed by Popper (1959), Lakatos (1968, 1971) posits that a research program is deemed progressive if it successfully establishes a core that withstands most attempts at disconfirmation, though not necessarily all.[6] Therefore, the role of confirmatory research in generating the core is not as an *absolute* singular step but rather involves developing scientific consensus through multiple confirmatory steps taken by a series of studies conducted within the research program. The role of exploratory research is pivotal in this context, as it contributes to the creation and extension of the core through a series of auxiliary hypotheses. These auxiliary hypotheses serve as inputs to the confirmatory research, allowing for a dynamic interplay between exploratory and confirmatory modes within the research program (Lieberman, 2020).

When exploring a novel phenomenon with a weak theoretical foundation, researchers typically initiate the research program with an inductive approach (Wold, 1982). Conversely, in the case of an established research program with an existing theoretical core, researchers typically rely on a deductive approach as the starting point. These entry points play a crucial role in shaping the trajectory of a research program over time. Figure 1 elucidates the complementary and cyclical nature of these approaches with relevant starting points, illustrating how the exploratory-first mixed EP mode and confirmatory-first mixed EP mode can iterate within themselves and in conjunction with each other to formulate a robust research program for EP

[6] Naïve falsificationism posits that a theory or hypothesis can be decisively rejected on the basis of a single disconfirmatory evidence (Kent, 2006).

theories. More importantly, the figure illustrates the guardrails associated with each stage with a focus on enhancing the replicability of the research program by adopting open science practices.

The exploratory EP mode indicated on the left side of Figure 1 provides a starting point for exploratory-first mixed EP mode research. As indicated by the first guardrail, the exploratory EP mode provides an opportunity for theory generation, engaging in intuitive and creative analysis of the data. This exploratory EP mode often leads researchers to engage in the illustrated inductive loop, indicating a theorizing process that is common in the minds of scholars, but rarely included in academic publications due to its lack of grounding in existing theory or data. The guardrail of preregistration to inform exploratory intent seeks to legitimize this inductive loop and its inclusion in published research findings. The third guardrail notes that exploratory EP mode research should be clearly labeled as exploratory, also including the rationale behind the decision made to engage in this mode of research. The final guardrail necessitates transparent reporting of achieved improvements in explanatory power and predictive accuracy including sample description and algorithm settings. The inclusion of such reports further strengthens the findings in exploratory EP mode studies. Research that begins in exploratory EP mode can then follow the exploratory-first mixed EP mode cycle indicated in Figure 1, where the theory generated in the exploratory EP mode is then tested in the confirmatory EP mode.

The confirmatory EP mode indicated on the right side of Figure 1 houses the majority of the published MIS research and provides an opportunity for theory confirmation. In the confirmatory EP mode, researchers may work in the illustrated deductive loop of confirming hypotheses made before analyzing the data set. As indicated by the first guardrail, confirmatory EP mode requires a priori hypotheses and a one-shot analysis of the data. The second guardrail is the preregistration of the confirmatory hypotheses, including a detailed pre-analysis plan. As indicated by the third guardrail, a separate sample from the one used in the exploratory EP mode must be used, as well as in each subsequent confirmatory analysis. Preregistration and the use of a separate sample are keys to maintaining the integrity of the illustrated deductive loop. The final guardrail calls for transparent confirmation of explanatory power and predictive accuracy, highlighting the value of the confirmatory EP mode findings, regardless of whether they provide confirmation for the stated hypotheses or not. This process of theory confirmation can then lead to further theory generation, highlighting the cyclical nature of these modes. This cyclical nature indicates that, regardless of whether a research project begins at the starting point of exploratory-first or confirmatory-first mixed EP mode, there is potential to broaden the scope of the research program by engaging in this iterative cycle.

The EP-mixed cycle identifies and accounts for both theory generation and testing stages in a research program, including exploratory research that is not typically reported in published research. Journals typically prefer to publish theoretically motivated confirmatory research, which compels researchers to present their exploratory research as if it was conducted using hypothetico-deductive method (Petter & Hadavi, 2021). The EP-mixed cycle and accompanying guardrails suggest that even if exploratory research is ultimately not published, preregistering the intent to conduct exploratory research and the transparent documentation of results achieved using the exploratory EP mode leaves behind a visible and traceable footprint for subsequent research to follow and enhance replicability of the work. This can help bring exploratory research into the open, as it is in other experimental natural sciences, giving proper credit to its valuable role in theory generation (Adler et al., 2023). This aspect is assuming even greater importance in the era of big data research, where MIS researchers confront novel questions and untested effects with a weak theory (Johnson et al., 2019; Maass et al., 2018).

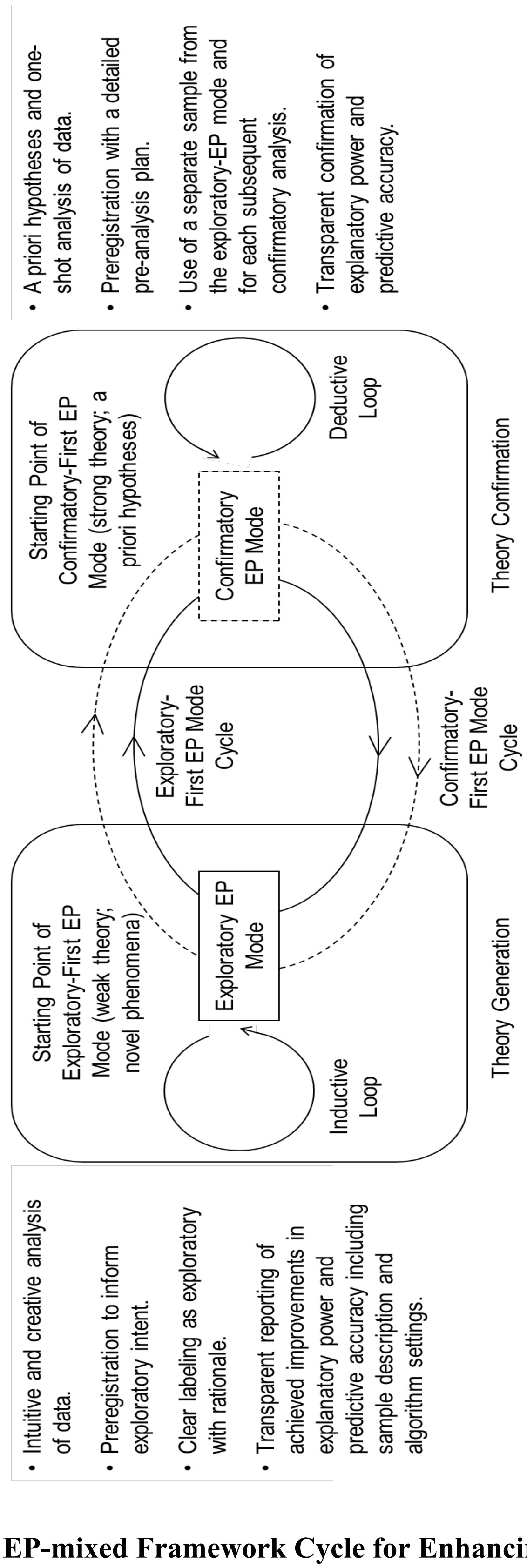

Figure 1: The EP-mixed Framework Cycle for Enhancing Replicability

6 Operationalizing the EP-mixed framework for replicable research using PLS-SEM

There is widespread interest in the social science community to integrate both explanatory and predictive assessments to enhance the replicability of findings. On one hand, computational social scientists urgently advocate for enhancing the explanatory capabilities of their machine-learning models (e.g., Athey, 2017; Hofman et al., 2017; Hofman et al., 2021; Wager & Athey, 2018). On the other hand, behavioral science researchers are eager to develop the predictive abilities of their explanatory models (e.g., Douglas, 2009; Shmueli & Koppius, 2011; Ward et al., 2010; Yarkoni & Westfall, 2017). The growing awareness of the weaknesses in external validity and replicability of our models has intensified the interest in concurrently employing novel explanatory and predictive analytic methods to establish robust EP theories.

The EP-mixed framework conceptually underpins the development of replicable EP theories in MIS research. To leverage the framework, we discuss analytical methods available in PLS-SEM that support empirical assessments required to develop EP theories in MIS research. PLS-SEM is known for its robust explanatory and predictive capabilities (e.g., Cepeda Carrión et al., 2016; Richter et al., 2022; Sharma et al., 2023) as well as its growing popularity in management disciplines for conducting both confirmatory and exploratory research (e.g., Benitez et al., 2020; Richter et al., 2016a; Sarstedt et al., 2022a). Furthermore, Adler et al. (2023) provide a comprehensive PLS-SEM specific preregistration template that could be utilized by researchers to implement the EP-mixed framework and enhance the replicability of their studies. These facets make PLS-SEM an attractive technique for developing and validating robust, replicable EP theories in MIS research. Table 4 presents the PLS-SEM methods that support the E, P, and, when combined in a stepwise fashion, the EP analytical mode. We discuss each in turn below.

Explanatory (E) mode methods (in-sample)		
Method	**Notes**	**Literature**
Path coefficients and effect sizes	Report the path coefficient estimates and effect sizes to quantify the strength of the relationships and their significance (e.g., using bootstrapping confidence intervals).	Chin (1998); Hair et al. (2022, Chap. 6); Henseler et al. (2009); Tenenhaus et al. (2005)
Explained Variance (R^2)	The higher the explained variance, the better is the in-sample explanation. However, one should not solely focus on maximizing the R^2 as this practice may lead to models that overfit the data.	Hair et al. (2019); McNeish (2015); Sharma et al. (2019); Shmueli (2010); Shmueli and Koppius (2011)
Model fit metrics	Utilize model fit metrics (e.g., SRMR) and bootstrap-based tests for model fit, while also considering the nuances regarding their appropriate application in PLS-SEM.	Hair et al. (2019); Ringle et al. (2023); Lohmöller (1989, Chaps. 2 and 5); Schuberth et al. (2023)
Information criteria	Focus on a key target construct of interest in the model and use the BIC/GM value to decide if the in-sample explanation has improved in an alternative model. The lower the BIC/GM value, the better the explanation.	Sharma et al. (2019); Sharma et al. (2021)
Akaike weights	For a key target construct of interest in the model and use the Akaike weights to decide if the in-sample explanation has been improved in an alternative model. Select the model based on the highest Akaike weight, which is a percentage value representing the relative likelihood.	Danks et al. (2020); Rigdon et al. (2023)
Model specification search	For in-sample criteria such as R^2 (and similar in-sample criteria) specification search algorithms in PLS-SEM support researchers by revealing alternative models with improved explanatory capabilities.	Cho et al. (2022a); Marcoulides and Drezner (2001); Marcoulides and Drezner (2003); Marcoulides et al. (1998)
Predictive (P) mode methods (out-of-sample)		
Method	**Notes**	**Literature**
Predictive power assessment using $PLS_{predict}$	Focus on a key target construct of interest in the model. The $PLS_{predict}$ results allow to determine if the PLS-SEM model has better predictive capabilities compared to benchmark methods (e.g., indicator average or linear model estimation). Based on these results across all indictors of the construct, a researcher can decide if the PLS-SEM model has sufficient predictive power.	Shmueli et al. (2016); Shmueli et al. (2019)
Predictive benchmarking assessment using the Cross-validated Predictive Ability Test (CVPAT)	For one or more target constructs in the model, CVPAT allows benchmarking whether the PLS-SEM model has significantly better predictive accuracy than naïve benchmarks such as indictor averages or a linear model estimation.	Liengaard et al. (2021); Sharma et al. (2023)
Predictive model comparison using CVPAT	For one or more target constructs in the model, CVPAT allows testing whether an alternative model has significantly better predictive accuracy than an established model.	Liengaard et al. (2021); Sharma et al. (2023)

Table 4: Analytical methods in PLS-SEM supporting the E and the P mode assessments

6.1 In-sample explanation methods in PLS-SEM

When the focus is on assessing the explanatory power of models, researchers typically assess path coefficient estimates and effect sizes to quantify the strength of the relationships and their statistical significance (e.g., using bootstrapping confidence intervals). Further, they use the R^2 values to assess the in-sample explained variance. The higher the value, the higher is the in-sample explanatory power. However, a sole focus on improving the R^2 value can cause model overfitting to the idiosyncratic noise in the data, with undesirable effects on the out-of-sample predictive capabilities of the model. This is also why researchers frequently use the R^2 to detect overfitting. As Hair et al. (2019, p. 11) note, "*when measuring a concept that is inherently predictable, such as physical processes, R^2 values of 0.90 might be plausible. Similar R^2 value levels in a model that predicts human attitudes, perceptions and intentions likely indicate an overfit.*"

The overfitting tendency proves particularly problematic in model comparisons where the R^2 almost always favors a more complex model (i.e., the model with a greater number of explanatory constructs). A more promising approach for selecting models is the use of Information Criteria. Sharma et al. (2019) and Sharma et al. (2021) show that the Bayesian information criterion (BIC; Schwarz, 1978) and the Geweke-Meese criterion (GM; Geweke & Meese, 1981) are particularly suitable for this purpose. The lower the BIC or GM value of a model's key target construct, the better the in-sample explanatory capability of the model. However, the differences in BIC or GM values between the models are sometimes relatively small, making it difficult to assess whether the alternative model really represents a substantial improvement for in-sample explanation. For this reason, Danks et al. (2020) and Rigdon et al. (2023) suggest using Akaike weights in PLS-SEM to document the uncertainty involved in model selection (see also Sarstedt & Moisescu, 2023).

Researchers have also proposed model fit metrics, which follow a psychometric perspective in that they assess a model's adequacy on the grounds of a (mis)match between the observed and model-implied covariance matrices—as opposed to the R^2, which considers model fit from an explanatory power perspective, thereby following an econometric tradition. Corresponding criteria include the standardized root mean square residual (SRMR) and the bootstrap-based tests for model fit (Schuberth et al., 2023).

Finally, it is worth mentioning that model specification search algorithms can help reveal alternative models that improve in-sample explanatory capabilities in the exploratory mode. Examples of approaches are the tabu search algorithm (Marcoulides et al., 1998), the genetic algorithm (Marcoulides & Drezner, 2001), and the ant colony optimization algorithm (Marcoulides & Drezner, 2003), which have been used for model specification search in an SEM—albeit not PLS—context. In addition, Cohen's method to explore path directionality (Callaghan et al., 2007; Cohen et al., 1994), the fuzzy-set qualitative comparative analysis (fsQCA) (Mikalef & Pateli, 2017; Ragin, 2008; Rasoolimanesh et al., 2021; Yan et al., 2023), and the necessary condition analysis (NCA; Dul, 2016; Hauff et al., 2024; Richter et al., 2022; Richter et al., 2023) may support researchers by revealing alternative model specifications. For example, the combined use of PLS-SEM and NCA may disclose specific relations as neither sufficient nor necessary, which researchers may delete to safeguard model parsimony—provided that theoretical considerations support such a step.

6.2 Out-of-sample prediction methods in PLS-SEM

Establishing a model's predictive power is of fundamental concern when working with regression-based methods, including PLS-SEM (Hofman et al., 2017; Hofman et al., 2021; Shmueli, 2010; Shmueli & Koppius, 2011). Management researchers often aim to provide practically relevant managerial recommendations, which are predictive by nature (Guenther et al., 2023; Hair, 2021; Magno et al., 2022; Sarstedt & Danks, 2022). Consequently, it is also important to assess the predictive power of models (Hair, 2021; Hair et al., 2022). The $\text{PLS}_{\text{predict}}$ procedure (Shmueli et al., 2016; Shmueli et al., 2019) and the cross-validated predictive ability test (CVPAT; Liengaard et al., 2021; Sharma et al., 2023) allow out-of-sample predictive power assessments of PLS path models. $\text{PLS}_{\text{predict}}$ focuses on a key target construct of interest in the model and its indicators. By using the principle of *k*-fold cross-validation, $\text{PLS}_{\text{predict}}$ compares predictive capability of the model for each indicator with prediction benchmarks (e.g., indicator averages or linear model estimation). Based on how often the PLS-SEM results beat the prediction benchmark for the set of a target construct's set of indicators, researchers can decide if the PLS-SEM estimation of the model has predictive power. However, just counting the number of indicators in favor of PLS-SEM is somewhat arbitrary, especially in (close to) split situations (e.g., when PLS-SEM provides better predictions for two indicators of the construct, while for the remaining two indicators the benchmark method provides better predictions). Thus, a limitation of $PLS_{predict}$ is that it relies on a heuristic approach for decision-making by depending on indicator counts rather than providing a comprehensive assessment of the construct or implementing a statistical test. To overcome this issue, Liengaard et al. (2021) and Sharma et al. (2023) introduced CVPAT to assess whether PLS-SEM model estimations beat naïve prediction benchmarks by offering statistically significant better predictive capabilities. This test is flexible and can focus on a single target construct of interest or a set of target constructs simultaneously (i.e., the prediction results of all selected constructs are included simultaneously for computing the test statistic). The CVPAT results facilitate both the predictive model assessment of specific models and the predictive model comparisons of alternative models.

Both in-sample and out-of-sample methods in PLS-SEM (Table 4) provide valuable support for researchers engaged in exploratory, confirmatory, and mixed EP mode analyses. To summarize, in the exploratory EP mode, a combination of in-sample model fit comparison metrics and the CVPAT can be employed to evaluate whether the introduction of a variable, path, moderator, or mediator enhances the model (refer to Table 2). Similar considerations apply to the confirmatory EP mode, which necessitates the formulation of robust a priori theory-based hypotheses. In the context of EP-mixed mode analyses, involving both exploration and confirmation, both in-sample model fit and out-of-sample predictive power must be considered simultaneously.

6.3 Further considerations and methodological extensions in PLS-SEM

Our discussion above opens avenues for further considerations and methodological extensions in PLS-SEM to aid the creation and validation of replicable EP theories. To enhance the explanatory aspects of their modeling, researchers can use latent class segmentation techniques to uncover the heterogeneity in their data and to form segments, which show significantly different coefficients in their segment-specific solutions and collectively provide (e.g., by their weighted average) a higher in-sample explanation across segments than the aggregate dataset model estimation results. For instance, various segmentation methods have been developed for this purpose that encompass finite mixture partial least squares (FIMIX-PLS; Hahn et al., 2002; Sarstedt et al., 2011; Sarstedt et al., 2022b), distance-based segmentation (Esposito Vinzi et al., 2008) including prediction-oriented segmentation (PLS-POS; Becker et al., 2013), decision trees

(Sánchez, 2009), evolutionary algorithms for segmentation (Ringle et al., 2014; Ringle et al., 2013), and iterative reweighted regression segmentation (PLS-IRRS; Schlittgen et al., 2016). The objective criterion of these approaches usually aims at maximizing in-sample explanation per segment (e.g., based on the segment-specific R^2 values). The final decision about the number of segments is based on segment retention criteria, the segment-specific differences of coefficients in the model, and the relative sample sizes of segments. The combined use of FIMIX-PLS and PLS-POS allows researcher to decide on a suitable number of segments based on the FIMIX-PLS segment retention criteria (e.g., Hair et al., 2016; Matthews et al., 2016) and the final segmentation into dichotomous groups based on PLS-POS (Hair et al., 2024, Chapter 6; Sarstedt et al., 2017).

Furthermore, future research could consider methodological improvements for predictive model specification search using out-of-sample prediction assessment criteria. This would support researchers in identifying alternative models with improved predictive capabilities. As an alternative to PLS-SEM (Cho et al., 2023; Cho et al., 2022b; Hwang et al., 2020), researchers can also utilize the component-based structural equation modeling method GSCA (Hwang et al., 2023; Hwang & Takane, 2004, 2014), which supports model selection (Cho et al., 2019) and model specification search (Cho et al., 2022a). Such methodological advancements could pave the way for the development of novel segmentation techniques that utilize out-of-sample prediction as an objective criterion. In addition, combining modeling methods such as GSCA and PLS-SEM with selected machine learning algorithms (e.g., Cho & Hwang, 2023; Richter & Tudoran, 2024) enables researchers to improve theoretical insight and predictive accuracy in MIS research and other disciplines.

7 Discussion and recommendations

In response to the growing concerns regarding the replication crisis in broader behavioral research, the MIS discipline has placed renewed emphasis on replicability. Initiatives like the publication of *AIS Transactions on Replication Research* and the *IS Replication Project* aim to address the replication crisis and establish replication assurance (Dennis et al., 2020). The burgeoning open science literature has provided valuable insights on distinguishing between exploratory and confirmatory modes and incorporating open science practices, such as preregistration (Adler et al., 2023; Nosek et al., 2018). However, despite these advancements, there are currently no established guidelines on effectively employing exploratory and confirmatory modes simultaneously in a mixed mode approach across various stages of the theory development cycle.

Additionally, while behavioral management research has recognized the utility of predictive modeling for theory development, proposing various predictive-inference tools (Cho et al., 2019; Liengaard et al., 2021; Shmueli et al., 2016), little attention has been paid to their critical role in enhancing replicability in MIS research (Shmueli & Koppius, 2011; Yarkoni & Westfall, 2017) and how they can be synergistically employed with explanatory modeling to create theories that both explain and predict—a critical goal in MIS research (Gregor, 2006). If MIS research seeks to build theories that both explain and predict, then it needs to adopt methods that assess not only how well the models explain the observed data but also assess how well they predict unobserved data.

Most significantly, to date, no practical frameworks have existed to guide researchers in organizing their thought processes for conducting replicable research by fully harnessing both exploratory and confirmatory modes, while aiming to construct explanatory and predictive theories. Without such a framework, researchers may continue falling into the trap of conflating exploratory with confirmatory research and relying solely on explanatory modeling, while overlooking the predictive modeling tools available at their disposal. We present a comprehensive framework that leverages predictive analytic techniques to enhance the replicability of behavioral theories in the MIS discipline and beyond.

The EP-mixed framework empowers researchers to select the optimal combination of theorization and analytical modes, contingent upon two critical considerations: (1) the state of theory and existing research in the field of study (exploratory or confirmatory), and (2) the goals of the study (explanation or prediction). The framework provides several benefits: First, it provides researchers with the capability to select an approach that aligns with the goal of enhancing the replicability of their work. Second, it encourages researchers to articulate the objectives of their research in a clear and concise manner, facilitating a better understanding for reviewers and journal editors and aiding in the application of appropriate judgment criteria. Third, despite the prevalent preference in management journals for publishing articles using confirmatory approaches (Adler et al., 2023), our framework establishes a foundation for the amalgamation of exploratory and confirmatory methods. This, in turn, lends legitimacy to innovative exploratory research addressing cutting-edge phenomena. Fourth, our framework positions predictive modeling at the forefront of theory building and clarifies how it can be integrated with explanatory modeling—a response to the calls made by various researchers (e.g., Hofman et al., 2017; Shmueli & Koppius, 2011; Ward et al., 2010; Yarkoni & Westfall, 2017). Finally, our framework opens further avenues of research by helping improve the predictive abilities and replicability of theories of novel MIS phenomena.

The EP-mixed framework comprises four theorization modes: confirmatory, exploratory, confirmatory-first mixed, and exploratory-first mixed, facilitating a flexible approach to research design. Within each mode, the EP-mixed framework provides clear prescriptions for analytical methods, guidelines, and metrics to assess both explanatory power and predictive accuracy. Furthermore, the framework suggests how these modes may be utilized to create new research programs or extend existing ones focusing on replicable EP theories.

It is crucial to emphasize that due to its strict hypothetico-deductive focus on robust theory testing, confirmatory research is less conducive to fostering innovation (Tukey, 1980). This constraint arises because the confirmatory mode is, by definition, *theory first*, and existing theories cannot always be expected to generate novel, unanticipated hypotheses (Guzzo et al., 2022). This limitation in the confirmatory mode becomes even more apparent in applied disciplines like MIS, which deal with emerging phenomena and complex data involving many untested variables and unknown effects. In contrast, the exploratory mode allows analytical flexibility with the potential to lead to newer insights into novel, unexplored phenomena (Adler et al., 2023). Nonetheless, the exploratory mode's emphasis on *sensitivity*—the ability to detect effects—renders it susceptible to false positives. This is where the confirmatory mode excels, given its emphasis on *specificity*—the ability to exclude non-effects—through strict adherence to an a priori plan and more conservative tests (Kimmelman et al., 2014).

The EP-mixed framework provides guidance to authors on utilizing exploratory and confirmatory modes concurrently in a mixed mode and striking a judicious balance between the inherent strengths and weaknesses of confirmatory and exploratory approaches by seamlessly integrating the two. The confirmatory-first mixed EP mode is suitable for extending existing EP theories in an incremental fashion using post hoc tests such as moderators, mediators, levels of analysis, or other contextual differences. It is important to note, however, that such post hoc findings are considered tentative until they are verified using future confirmatory studies.

Conversely, when investigating entirely novel phenomena with weak or nonexistent theory, the exploratory-first mixed EP mode becomes the preferable choice. The exploratory-first mixed EP mode represents the most innovative yet robust approach, providing freedom in the initial exploratory stage followed by a rigorous confirmatory stage to instill confidence in the findings using a separate sample. Essentially, the confirmatory stage in the exploratory-first mixed EP mode serves to provide replication assurance. When engaged in the initial exploratory mode, researchers should clearly label analyses as exploratory, and provide necessary details for replication of results. In the subsequent confirmatory mode, researchers should use a separate sample and adhere to a preregistered plan and one-shot analyses to maintain alignment with a priori hypothesizing, minimizing the risk of *p*-hacking or HARKing.

The process of theory generation followed by theory testing (or vice versa) using the EP-mixed framework can, in principle, be applied iteratively to conduct a research program as long as the guidelines for each stage are appropriately followed (refer to Figure 1). Such a program would concurrently embed rigorous testing within the process of discovery. The proposed EP-mixed framework provides a roadmap for researchers to navigate the complexities of novel theory design in MIS research, ensuring that theories are rigorously tested and generalizable. The framework emphasizes the valuable and distinct contributions of both exploratory and confirmatory research, underscores the importance of clearly distinguishing between them, and suggests ways in which they can be effectively utilized together. Each of the four theorization modes defined by the EP-mixed framework, when conducted according to the referenced guidelines, has unique potential to contribute to both theory and practice in MIS research. By integrating predictive analytic techniques and emphasizing the importance of pre-analysis plans (preregistration) and transparent reporting, the framework seeks to address the replication crisis in behavioral research and paves the way for more robust and generalizable theories in the MIS and beyond. With its clear guidelines and tools, the EP-mixed framework holds potential to enhance the rigor of MIS research.

We believe that the replication crisis in behavioral research represents an opportunity for advancements in MIS research. The EP-mixed framework provides a means through which to move forward in this direction to support researchers in their decisive actions to ensure the replicability of MIS models. Moreover, this comprehensive approach aligns with the longstanding pursuit of MIS to enhance theoretical and practical relevance. Recognizing the priority placed on theoretically relevant models, researchers should transparently communicate instances where models may fall short of achieving practical relevance. This commitment to transparency contributes to the ongoing quest for models that not only advance theoretical understanding but also exhibit real-world applicability. Based on our discussion we provide certain recommendations for authors and researchers leveraging the EP-mixed framework.

Recommendation 1: Choose and apply the theorization mode wisely. Swedberg (2020) emphasizes the critical role of exploratory research, describing it as the *soul* of high-quality social science research. He asserts that without the ambition to contribute novel insights using exploratory research, the progress of research would stagnate, since non-exploratory (confirmatory) research, by its very definition, can only result in the repetition of existing knowledge. Traditionally, exploratory research has been associated with qualitative methodologies such as case studies or grounded theory (Stebbins, 2001). However, in the era of big data, exploratory MIS research is increasingly reliant on quantitative methodologies (Johnson et al., 2019; Maass et al., 2018).

The exploratory mode provides a pathway for MIS researchers to investigate cutting-edge socio-technical issues and conduct innovative research. In this paradigm, the role of the confirmatory mode is to furnish more definitive evidence of the findings. The primary challenge for MIS researchers lies in effectively integrating these two modes and to avoid contributing to a replication crisis through indiscriminate mixing (Nosek et al., 2018).

To address this challenge, researchers should articulate a compelling rationale for their selection among the four theorization modes presented in the EP-mixed framework—confirmatory, exploratory, confirmatory-first, and exploratory-first modes. This decision should be grounded in an assessment of the current state of theoretical development related to the phenomenon of interest, guided by questions such as: "Is the socio-technical phenomenon under investigation genuinely novel?" and "To what extent does existing theory or literature offer guidance in explaining the phenomenon?" The answers to these questions will predominantly influence the choice of theorization mode. Subsequently, the EP-mixed framework provides guidance on the associated guardrails for applying each mode in a robust fashion.

Recommendation 2: Choose and apply the analytical mode wisely. Researchers should articulate a compelling rationale when choosing among the analytical modes outlined within the EP-mixed framework—identified as E, P, and EP—and utilize appropriate metrics to substantiate their decision. This rationale should be contingent upon the specific objectives of the theory and the associated inferences. If the primary goal is to explain the underlying mechanisms driving the observed data, researchers should opt for in-sample metrics conducive to the E mode. Conversely, when the aim is to provide prospective inferences or prescriptive guidance to decision-makers, researchers should employ out-of-sample metrics aligning with the P mode. Without such transparency, the implicit assumptions underlying each approach could lead to different conclusions (Sarstedt et al., 2024).

Moreover, it is crucial for researchers to appreciate that in-sample assessments may not robustly support out-of-sample inferences, and vice versa (Shmueli & Koppius, 2011). In the development of theories encompassing both explanatory and predictive dimensions (i.e., EP theories), researchers should rigorously assess models for their explanatory efficacy using in-sample methods, while also evaluating their predictive capacities through out-of-sample approaches. This holds particular significance for MIS research where the majority of theories aspire to not only explain but also to predict (Gregor, 2006). For a more comprehensive understanding and guidance on substantiating these decisions, readers are referred to seminal works by Shmueli (2010), Shmueli and Koppius (2011), and Yarkoni and Westfall (2017).

Recommendation 3: Choose and utilize a replication sample wisely. The use of a separate replication sample is necessary for separating exploratory and confirmatory modes in a study (Nosek et al., 2018). This aspect is most relevant for the exploratory-first mode because the sample that was used to generate the hypotheses should not be used to test (confirm) the same hypotheses. The choice of the confirmation sample, and whether a new data collection effort is required, depends on a theoretically motivated and well-defined target population to which the study's inferences are intended for extrapolation and the level of external validity sought by the researcher.

A successful replication of results when the sample used for confirmation comes from the same target population as the sample used for exploration provides evidence of context-specific *generalizability* of effects. In contrast, a successful replication of results in a different target population is evidence of the *transportability* of effects and provides greater confidence in the external validity of the study across contexts (Findley et al., 2021).

In the former scenario, a researcher may opt for either collecting an entirely new sample from the same population or employing strategies such as a random 50-50 split of the original data. However, it should be noted that the random split-sample approach has well-documented limitations, including reduced statistical power and suboptimal predictive performance due to the halving of the original sample size. Specifically, the confirmatory sample should be sufficiently large to minimize the effect of random variation and false positives (Kimmelman et al., 2014). Hence, unless the original sample size is substantial, a 50-50 split-sample method is not recommended (Kimmelman et al., 2014; Steyerberg & Harrell, 2016).

In the latter scenario, a separate data collection effort from a distinct target population is obviously required to provide evidence of transportability. An illustrative example of this approach is the study by Aljukhadar and Senecal (2016) that assessed the differential moderating roles of website expertise and e-commerce expertise on the ease of use perceptions and acceptance of a website among English-speaking users. Subsequently, they replicated the effects of their study using a separate sample of French-speaking users, thereby providing evidence of the transportability of effects.

Recommendation 4: Utilize open science practices such as preregistrations whenever feasible. Researchers are encouraged to embrace the open science practices of preregistration to enhance transparency in their analysis, results reporting, and conclusions, and open sharing of data whenever feasible.[7] In the specific context of PLS-SEM, Adler et al. (2023) provide a comprehensive preregistration template that can be utilized for this purpose. One may wonder how detailed a preregistration should be. There is no "one size fits all" recommendation as the level of detail may vary depending on the context of the study and the constraints faced by the researcher.

In the confirmation mode, a basic preregistration may include information regarding the specific research questions and a priori hypotheses selected for examination, as well as details pertaining to the sample population. This might include general information regarding sample selection procedures (e.g., sample size, inclusion/exclusion criteria), survey items, and the statistical methodologies to be used (Claesen et al., 2021). More comprehensive preregistrations may incorporate finer details such as specific algorithm configurations and settings, precise sample selection strategies, variables and operationalizations, and specific statistical tests

[7] We acknowledge that open sharing of data, in particular, may not always align with researchers' incentives. Kwon and Motohashi (2021) discuss this issue in more detail by exploring the short versus long term benefits of open data sharing for researchers.

including robustness checks and endogeneity tests (Adler et al., 2023). Clearly, the more detailed the preregistration and stricter the adherence to the pre-analysis plan, the higher the confidence in the confirmatory character of the study (Nosek et al., 2018).

In contrast, a detailed preregistration for an exploratory study is not practical. Instead, researchers operating in the exploratory mode may choose to utilize preregistration as a proactive step to signal their intention to engage in exploratory research in advance (Fife & Rodgers, 2022; Szollosi et al., 2020), or they may choose to bypass preregistration entirely. What is essential in the exploratory mode, however, is transparent communication of findings and the methodologies employed. This entails detailing factors such as algorithm settings, operationalizations, tests conducted, treatment of outliers and missing values, as well as other data handling procedures used to achieve the results.

It should be emphasized that, currently, preregistrations are not obligatory for MIS journals; hence, their use remains voluntary. We also acknowledge that there are many instances where preregistration is not feasible or even possible.[8] If a preregistration is not feasible in the confirmatory mode, researchers should maintain a strict adherence to an a priori plan with one-shot analyses and transparent reporting of results.

Recommendation 5: Report deviations (if any) from a preregistered plan in the confirmatory mode. Thus far, we have suggested that a confirmatory study relies on one-shot analyses based on a rigid pre-analysis plan which is ideally preregistered. A one-shot analysis assumes that the researcher has *calibrated the instruments* sufficiently prior to collecting data, or in the case of secondary data, prior to analyzing it. Gaus et al. (2015, p. 3) note that, "*Strictly speaking, to facilitate a confirmative interpretation only one chance to identify a significant result is allowed, i.e. only one test is permitted*."

However, during the course of sample collection and analysis, a strict adherence to the plan is not always possible or feasible due to a variety of unanticipated methodological

[8] It is crucial to recognize that preregistrations may not always be feasible, and their application may disproportionately favor confirmatory studies, particularly in applied disciplines such as MIS (Guzzo et al., 2022). This poses a challenge, as the strict adherence to open science practices, including preregistration, has faced criticism for potentially undervaluing the significance of innovative exploratory research. Striking a balance between promoting transparency and acknowledging the inherent flexibility required in certain disciplines is essential to ensure the continued appreciation and encouragement of both confirmatory and exploratory approaches. For example, Guzzo et al. (2022, p. 499) note: "*Open science practices clearly emphasize the hypothetico-deductive model of theory testing and creation and, consequently, tilt science–practice disciplines away from alternative ways of theorizing that are especially powerful in applied research*." However, the EP-mixed framework provides a way forward by giving the due credit to valuable exploratory research in a mixed mode format.

considerations affecting sample size, exclusion criteria (e.g., outliers or heteroscedasticity), or issues with statistical analysis (Claesen et al., 2021). Open science advocates refer to such methodological adjustments as *deviations* from the pre-analysis plan that should be transparently reported in the study. If the authors and reviewers consider such deviations as routine, minor, and not prone to publication bias, then a study may retain its confirmatory character, as long as the deviations are reported and justified.

For example, a researcher may have planned to conduct a single MANOVA analysis but instead switches to conducting two separate ANOVAs based on a weak correlation between the two dependent variables. In this case, a clear rationale should be provided for why the switch is being made (perhaps conducting two separate ANOVAs allows a simpler and more intuitive way to visualize the slopes). In contrast, deviations from the plan that are deemed significant (e.g., significant changes in the variables or their operationalizations, model, or a change in sampling strategy) could impinge the purported confirmatory nature of the study, and researchers should consider relabeling their study as exploratory (Nosek et al., 2018).

Finally, journal editors and reviewers also have a crucial role in addressing the ongoing replication crisis. A fundamental shift is needed, emphasizing an openness to mixed mode studies. Many management journal reviewers and editors have been hesitant to publish exploratory research, often with valid concerns. This reluctance stems from worries about researchers' analytical liberties, data dredging, transparency issues, and weak theoretical support associated with exploratory studies (Adler et al., 2023). Consequently, a substantial proportion of published research adheres to the hypothetico-deductive framework or at least claims to do so. Simultaneously, there is a persistent emphasis on novel theory development for cutting-edge phenomena not adequately explained by traditional theories—a scenario tailor-made for exploratory research. Unfortunately, this tension often creates a dilemma for researchers and contributes to the replication crisis.

Reviewers may also implicitly assume that in-sample explanatory assessments suffice to support out-of-sample prospective inferences, leading to the publication of overfit models that lack generalizability. The EP-mixed framework offers a solution by introducing two mixed EP modes that allow the integration of genuinely novel exploratory research with robust confirmatory research, each with separate guardrails, preserving their distinct character, requirements, and goals. The framework also facilitates and encourages incorporating predictive assessments when the focus is on creating EP theories.

A practical step toward breaking this cycle involves encouraging the publication of mixed mode studies. Transparent reports detailing exploratory analyses and the adoption of open science practices, such as preregistrations and data sharing, should be promoted wherever feasible. Registered reports serve as an alternative publication mechanism designed to provide assurance to researchers by focusing on the publication of results, even those that emerge from confirmatory tests that did not yield the expected outcomes (Adler et al., 2023). However, it is also important for reviewers note that strict adherence to preregistrations may not always be feasible or possible (Guzzo et al., 2022), and open data sharing may not always align well with researchers' incentives (Kwon & Motohashi, 2021). Unless required by journal policy, authors should not be penalized for failing to preregister or openly sharing the data. Instead, in such cases, reviewers should ensure that the guardrails distinguishing exploratory research from confirmatory research have been appropriately followed. Embracing a more inclusive mixed mode approach can create a dynamic academic environment conducive to advancing knowledge and resolving the replication crisis.

8 Conclusion

In this article, we sought to address the ontological and epistemological issues encountered by MIS researchers during novel theory development, particularly in light of recent concerns about the replicability of findings. We focused on two key factors contributing to the replication crisis in behavioral research: (1) the conflation of exploratory and confirmatory research and (2) the predominant emphasis on in-sample explanatory assessments with a corresponding neglect of out-of-sample predictive analytic assessments.

As a response to these challenges, we introduced the holistic EP-mixed framework, designed to assist researchers in navigating analytical approaches based on their research objectives and the nature of the theory under development. This framework encourages the integration of both exploratory and confirmatory research modes, allowing researchers to leverage the strengths of each. Additionally, it advocates for the simultaneous use of in-sample explanatory and out-of-sample predictive assessments, promoting the creation of robust and replicable EP theories that possess both theoretical and practical relevance.

While we used PLS-SEM as an exemplar method to operationalize the EP-mixed framework, the framework is generally adaptable and applicable across a diverse range of statistical methods. Researchers should choose statistical methods that best support their research objectives (e.g., Rigdon, 2024; Rigdon et al., 2017) or to adopt a multimethod approach to address methodological uncertainties (Sarstedt et al., 2024). By emphasizing mixed mode research, the EP-mixed framework provides a comprehensive approach to theory development, enhancing the replicability and robustness of MIS theories. Researchers are encouraged to utilize this framework in conjunction with various statistical methods to further advance the field and contribute to the development of EP theories with enduring value.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT 3.5, DeepL, and Google Gemini to copy edit and improve the language and readability of author-generated text. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Adler, S. J., Sharma, P. N., & Radomir, L. (2023). Toward Open Science in PLS-SEM: Assessing the State of the Art and Future Perspectives. *Journal of Business Research*, *169*, 114291.

Aljukhadar, M., & Senecal, S. (2016). The User Multifaceted Expertise: Divergent Effects of the Website versus E-commerce Expertise. *International Journal of Information Management*, *36*(3), 322-332.

Athey, S. (2017). Beyond Prediction: Using Big Data for Policy Problems. *Science*, *355*(6324), 483-485.

Bariff, M. L., & Ginzberg, M. J. (1982). MIS and the Behavioral Sciences: Research Patterns and Prescriptions. *SIGMIS Database*, *14*(1), 19–26.

Bawack, R. E., Bonhoure, E., Kamdjoug, J.-R. K., & Giannakis, M. (2023). How Social Media Live Streams Affect Online Buyers: A Uses and Gratifications Perspective. *International Journal of Information Management*, *70*, 102621.

Becker, J.-M., Rai, A., Ringle, C. M., & Völckner, F. (2013). Discovering Unobserved Heterogeneity in Structural Equation Models to Avert Validity Threats. *MIS Quarterly*, *37*(3), 665-694.

Benbasat, I., & Zmud, R. W. (1999). Empirical Research in Information Systems: The Practice of Relevance. *MIS Quarterly*, *23*(1), 3-16.

Benbasat, I., & Zmud, R. W. (2003). The Identity Crisis within the Is Discipline: Defining and Communicating the Discipline's Core Properties. *MIS Quarterly*, *27*(2), 183-194.

Benitez, J., Henseler, J., Castillo, A., & Schuberth, F. (2020). How to Perform and Report an Impactful Analysis Using Partial Least Squares: Guidelines for Confirmatory and Explanatory IS Research. *Information & Management*, *57*(2), 103168.

Blut, M., Chong, A. Y. L., Tsiga, Z., & Venkatesh, V. (2022). Meta-Analysis of The Unified Theory of Acceptance and Use of Technology (UTAUT): Challenging Its Validity and Charting A Research Agenda In The Red Ocean. *Journal of the Association for Information Systems*, *23*(1), 13-95.

Bogert, E., Schecter, A., & Watson, R. T. (2021). Preregistration of Information Systems Research. *Communications of the Association for Information Systems*, *49*(5), 110-125

Califf, C. B., Sarker, S., & Sarker, S. (2020). The Bright and Dark Sides of Technostress: A Mixed-Methods Study Involving Healthcare IT. *MIS Quarterly*, *44*(22), 809-856.

Callaghan, W., Wilson, B., Ringle, C. M., & Henseler, J. (2007). Exploring Causal Path Directionality for a Marketing Model: Using Cohen’s Path Method. PLS'07: The 5th International Symposium on PLS and Related Methods, Ås, Norway.

Carnap, R. (2009). Theory and Observation. In T. J. McGrew, M. Alspector-Kelly, & F. Allhoff (Eds.), *The Philosophy of Science: An Historical Anthology* (pp. 329-343). Wiley-Blackwell.

Cepeda Carrión, G., Henseler, J., Ringle, C. M., & Roldán, J. L. (2016). Prediction-Oriented Modeling in Business Research by Means of PLS Path Modeling. *Journal of Business Research*, *69*(10), 4545-4551.

Chen, S., Tran, K. T., Xia, Z., Waseem, D., Zhang, J. A., & Potdar, B. (2023). The Double-edged Effects of Data Privacy Practices on Customer Responses. *International Journal of Information Management*, *69*, 102600.

Chin, W. W. (1998). The Partial Least Squares Approach to Structural Equation Modeling. In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295-358). Lawrence Erlbaum.

Cho, G., & Hwang, H. (2023). Deep Learning Generalized Structured Component Analysis: An Interpretable Artificial Neural Network Model with Composite Indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, *forthcoming*.

Cho, G., Hwang, H., Sarstedt, M., & Ringle, C. M. (2022a). A Prediction-Oriented Specification Search Algorithm for Generalized Structured Component Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(4), 611-619.

Cho, G., Jung, K., & Hwang, H. (2019). Out-of-bag Prediction Error: A Cross Validation Index for Generalized Structured Component Analysis. *Multivariate Behavioral Research*, *54*(4), 505-513.

Cho, G., Lee, J., Hwang, H., Sarstedt, M., & Ringle, C. M. (2023). A Comparative Study of the Predictive Power of Component-based Approaches to Structural Equation Modeling. *European Journal of Marketing*, *57*(6), 1641-1661.

Cho, G., Sarstedt, M., & Hwang, H. (2022b). Comparison of Covariance Structure Analysis, Partial Least Squares Path Modeling and Generalized Structured Component Analysis in Factor- and Composite Models. . *British Journal of Mathematical and Statistical Psychology*, *75*(2), 220-251.

Claesen, A., Gomes, S., Tuerlinckx, F., & Vanpaemel, W. (2021). Comparing Dream to Reality: An Assessment of Adherence of the First Generation of Preregistered Studies. *Royal Society Open Science*, *8*(10), 211037.

Cohen, P. R., Hart, D. M., St. Amant, R., Ballesteros, L. A., & Carlson, A. (1994). Path Analysis Models of an Autonomous Agent in a Complex Environment. In P. Cheeseman & R. W. Oldford, *Selecting Models from Data* New York, NY.

Danks, N. P., Sharma, P. N., & Sarstedt, M. (2020). Model Selection Uncertainty and Multimodel Inference in Partial Least Squares Structural Equation Modeling (PLS-SEM). *Journal of Business Research*, *113*, 13-24.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319-340.

DeLone, W. H., & McLean, E. R. (1992). Information Systems Success: The Quest for the Dependent Variable. *Information Systems Research*, *3*(1), 60-95.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems*, *19*(4), 9-30.

Dennis, A., Brown, S. A., Wells, T. M., & Rai, A. (2020). Editor's Comments: Replication Crisis or Replication Reassurance: Results of the IS Replication Project. *MIS Quarterly*, *44*(3), iii-x.

Diamantopoulos, A. (1994). Modelling with LISREL: A Guide for the Uninitiated. *Journal of Marketing Management*, *10*(1-3), 105-136.

Douglas, H., & Magnus, P. D. (2013). State of the Field: Why Novel Prediction Matters. *Studies in History and Philosophy of Science Part A*, *44*(4), 580-589.

Douglas, H. E. (2009). Reintroducing Prediction to Explanation. *Philosophy of Science*, *76*(4), 444-463.

Dubin, R. (1969). *Theory Building: A Practical Guide to the Construction and Testing of Theoretical Models*. The Free Press.

Dubin, R. (1976). Theory Building in Applied Areas. In M. D. Dunnette (Ed.), *Handbook of Industrial and Organizational Psychology* (pp. 17-39). Rand-McNally.

Dul, J. (2016). Necessary Condition Analysis (NCA): Logic and Methodology of "Necessary but not Sufficient" Causality. *Organizational Research Methods*, *19*(1), 10-52.

Esposito Vinzi, V., Trinchera, L., Squillacciotti, S., & Tenenhaus, M. (2008). REBUS-PLS: A Response-Based Procedure for Detecting Unit Segments in PLS Path Modelling. *Applied Stochastic Models in Business and Industry*, *24*(5), 439-458.

Fife, D. A., & Rodgers, J. L. (2022). Understanding the Exploratory/Confirmatory Data Analysis Continuum: Moving Beyond the "Replication Crisis". *American Psychologist*, *77*(3), 453-466.

Findley, M. G., Kikuta, K., & Denly, M. (2021). External Validity. *Annual Review of Political Science*, *24*(Volume 24, 2021), 365-393.

Gallien, J., Graves, S. C., & Scheller-Wolf, A. (2016). Practice-Based Research in Operations Management: What It Is, Why Do It, Related Challenges, and How to Overcome Them. *Manufacturing & Service Operations Management*, *18*(1), 5-14.

Gaus, W., Mayer, B., & Muche, R. (2015). Interpretation of Statistical Significance - Exploratory Versus Confirmative Testing in Clinical Trials, Epidemiological Studies, Meta-Analyses and Toxicological Screening (Using Ginkgo biloba as an Example). *Journal of Clinical & Experimental Pharmacology*, *5*(4), 1000182.

Geweke, J., & Meese, R. (1981). Estimating Regression Models of Finite but Unknown Order. *Journal of Econometrics*, *16*(1), 162.

Gregor, S. (2006). The Nature of Theory in Information Systems. *MIS Quarterly*, *30*(3), 611-642.

Grover, V., & Lyytinen, K. (2023). The Pursuit of Innovative Theory in the Digital Age. *Journal of Information Technology*, *38*(1), 45-59.

Guenther, P., Guenther, M., Ringle, C. M., Zaefarian, G., & Cartwright, S. (2023). Improving PLS-SEM Use for Business Marketing Research. *Industrial Marketing Management*, *111*, 127-142.

Guzzo, R. A., Nalbantian, H. R., & Schneider, B. (2022). Open Science, Closed Doors: The Perils and Potential of Open Science for Research in Practice. *Industrial and Organizational Psychology*, *15*(4), 495-515.

Hahn, C., Johnson, M. D., Herrmann, A., & Huber, F. (2002). Capturing Customer Heterogeneity Using a Finite Mixture PLS Approach. *Schmalenbach Business Review*, *54*(3), 243-269.

Hair, J. F. (2021). Next Generation Prediction Metrics for Composite-based PLS-SEM. *Industrial Management & Data Systems*, *121*(1), 5-11.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (3 ed.). Sage.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to Use and How to Report the Results of PLS-SEM. *European Business Review*, *31*(1), 2-24.

Hair, J. F., & Sarstedt, M. (2021). Explanation Plus Prediction: The Logical Focus of Project Management Research. *Project Management Journal*, *52*(4), 319-322.

Hair, J. F., Sarstedt, M., Matthews, L., & Ringle, C. M. (2016). Identifying and Treating Unobserved Heterogeneity with FIMIX-PLS: Part I - Method. *European Business Review*, *28*(1), 63-76.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2024). *Advanced Issues in Partial Least Squares Structural Equation Modeling (PLS-SEM)* (2 ed.). Sage.

Hauff, S., Richter, N. F., Sarstedt, M., & Ringle, C. M. (2024). Importance and Performance in PLS-SEM and NCA: Introducing the Combined Importance-Performance Map Analysis (cIPMA). *Journal of Retailing and Consumer Services*, *78*, 103723.

Hempel, C. G. (1962). Two Basic Types of Scientific Explanation. In M. Curd & J. A. Cover (Eds.), *Philosophy of Science: The Central Issues* (pp. 685-694). Norton.

Hensel, P. G. (2021). Reproducibility and Replicability Crisis: How Management Compares to Psychology and Economics – A Systematic Review of Literature. *European Management Journal*, *39*(5), 577-594.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The Use of Partial Least Squares Path Modeling in International Marketing. In R. R. Sinkovics & P. N. Ghauri (Eds.), *Advances in International Marketing* (Vol. 20, pp. 277-320). Emerald

Hill, D. W., & Jones, Z. M. (2014). An Empirical Evaluation of Explanations for State Repression. *American Political Science Review*, *108*(3), 661-687.

Hofman, J. M., Sharma, A., & Watts, D. J. (2017). Prediction and Explanation in Social Systems. *Science*, *355*(6324), 486-488.

Hofman, J. M., Watts, D. J., Athey, S., Garip, F., Griffiths, T. L., Kleinberg, J., Margetts, H., Mullainathan, S., Salganik, M. J., Vazire, S., Vespignani, A., & Yarkoni, T. (2021). Integrating Explanation and Prediction in Computational Social Science. *Nature*, *595*(7866), 181-188.

Hu, H.-f., Moody, G. D., & Galletta, D. F. (2023). HARKing and P-Hacking: A Call for More Transparent Reporting of Studies in the Information Systems Field. *Communications of the Association for Information Systems*, *52*(41), 853-876.

Hult, G. T. M., Sharma, P. N., Morgeson, F. V., & Zhang, Y. (2019). Antecedents and Consequences of Customer Satisfaction: Do They Differ Across Online and Offline Purchases? *Journal of Retailing*, *95*(1), 10-23.

Hwang, H., Sarstedt, M., Cheah, J. H., & Ringle, C. M. (2020). A Concept Analysis of Methodological Research on Composite-based Structural Equation Modeling: Bridging PLSPM and GSCA. *Behaviormetrika*, *47*, 219–241.

Hwang, H., Sarstedt, M., Cho, G., Choo, H., & Ringle, C. M. (2023). A Primer on Integrated Generalized Structured Component Analysis. *European Business Review*, *35*(3), 261-284.

Hwang, H., & Takane, Y. (2004). Generalized Structured Component Analysis. *Psychometrika*, *69*(1), 81-99.

Hwang, H., & Takane, Y. (2014). *Generalized Structured Component Analysis: A Component-Based Approach to Structural Equation Modeling*. Chapman & Hall.

Jacobucci, R. (2022). A Critique of Using the Labels Confirmatory and Exploratory in Modern Psychological Research. *Front Psychol*, *13*, 1020770.

Jebb, A. T., Parrigon, S., & Woo, S. E. (2017). Exploratory Data Analysis as a Aoundation of Inductive Research. *Human Resource Management Review*, *27*(2), 265-276.

Jensen, T. I., Bryan, K., & Pedersen, L. H. (2023). Is There a Replication Crisis in Finance? *The Journal of Finance*, *78*(5), 2465-2518.

Jeyaraj, A., Dwivedi, Y. K., & Venkatesh, V. (2023). Intention in Information Systems Adoption and Use: Current State and Research Directions. *International Journal of Information Management*, *73*, 102680.

Johnson, S. L., Gray, P., & Sarker, S. (2019). Revisiting IS Research Practice in the Era of Big Data. *Information and Organization*, *29*(1), 41-56.

Jöreskog, K. G., & Wold, H. (1982). The ML and PLS Techniques for Modeling with Latent Variables: Historical and Comparative Aspects. In K. G. Jöreskog & H. Wold (Eds.), *Systems Under Indirect Observation, Part I* (pp. 263-270). North-Holland.

Kelley, K., & Preacher, K. J. (2012). On Effect Size. *Psychological Methods*, *17*(2), 137-152.

Kent, M. (2006). *The Oxford Dictionary of Sports Science & Medicine* (3 ed.). Oxford University Press.

Killeen, P. R. (2019). Predict, Control, and Replicate to Understand: How Statistics Can Foster the Fundamental Goals of Science. *Perspect Behav Sci*, *42*(1), 109-132.

Kimmelman, J., Mogil, J. S., & Dirnagl, U. (2014). Distinguishing between Exploratory and Confirmatory Preclinical Research Will Improve Translation. *PLOS Biology*, *12*(5), e1001863.

Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction Policy Problems. *American Economic Review*, *105*(5), 491-495.

Kumar, P., Sharma, S. K., & Dutot, V. (2023). Artificial Intelligence (AI)-enabled CRM Capability in Healthcare: The Impact on Service Innovation. *International Journal of Information Management*, *69*, 102598.

Kwon, S., & Motohashi, K. (2021). Incentive or Disincentive for Research Data Disclosure? A Large-scale Empirical Analysis and Implications for Open Science Policy. *International Journal of Information Management*, *60*, 102371.

Lakatos, I. (1968). Criticism and the Methodology of Scientific Research Programmes. *Proceedings of the Aristotelian Society*, *69*, 149-186.

Lakatos, I. (1971). History of Science and its Rational Reconstructions. In R. C. Buck & R. S. Cohen (Eds.), *PSA 1970: In Memory of Rudolf Carnap Proceedings of the 1970 Biennial Meeting Philosophy of Science Association* (pp. 91-136). Springer Netherlands.

Lakatos, I. (1976). Falsification and the Methodology of Scientific Research Programmes. In S. G. Harding (Ed.), *Can Theories be Refuted? Essays on the Duhem-Quine Thesis* (pp. 205-259). Springer Netherlands.

Lee, A. S. (1991). Integrating Positivist and Interpretive Approaches to Organizational Research. *Organization Science*, *2*(4), 342-365.

Lee, J. K., Park, J., Gregor, S., & Yoon, V. (2021). Axiomatic Theories and Improving the Relevance of Information Systems Research. *Information Systems Research*, *32*(1), 147-171.

Lieberman, E. (2020). Research Cycles. In C. Elman, J. Gerring, & J. Mahoney (Eds.), *The Production of Knowledge: Enhancing Progress in Social Science* (pp. 42-70). Cambridge University Press.

Liengaard, B. D., Sharma, P. N., Hult, G. T. M., Jensen, M. B., Sarstedt, M., Hair, J. F., & Ringle, C. M. (2021). Prediction: Coveted, Yet Forsaken? Introducing a Cross-validated Predictive Ability Test in Partial Least Squares Path Modeling. *Decision Sciences*, *52*(2), 362-392.

Lohmöller, J.-B. (1989). *Latent Variable Path Modeling with Partial Least Squares*. Physica.

Maass, W., Parsons, J., Purao, S., Storey, V. C., & 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), 1253-1273.

Magno, F., Cassia, F., & Ringle, C. M. (2022). A Brief Review of Partial Least Squares Structural Equation Modeling (PLS-SEM) Use in Quality Management Studies. *The TQM Journal*, *ahead-of-print*(ahead-of-print).

Maier, C., Thatcher, J. B., Grover, V., & Dwivedi, Y. K. (2023). Cross-sectional Research: A Critical Perspective, Use Cases, and Recommendations for IS Research. *International Journal of Information Management*, *70*, 102625.

Marcoulides, G. A., & Drezner, Z. (2001). Specification Searches in Structural Equation Modeling with a Genetic Algorithm. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced Structural Equation Modeling: New Developments and Techniques* (pp. 247-268). Lawrence Erlbaum.

Marcoulides, G. A., & Drezner, Z. (2003). Model Specification Searches Using Ant Colony Optimization Algorithms. *Structural Equation Modeling*, *10*(1), 154-164.

Marcoulides, G. A., Drezner, Z., & Schumacker, R. E. (1998). Model Specification Searches in Structural Equation Modeling Using Tabu Search. *Structural Equation Modeling: A Multidisciplinary Journal*, *5*(4), 365-376.

Matthews, L., Sarstedt, M., Hair, J. F., & Ringle, C. M. (2016). Identifying and Treating Unobserved Heterogeneity with FIMIX-PLS: Part II – A Case Study. *European Business Review*, *28*(2), 208-224.

McCain, K. (2022). *Understanding How Science Explains the World*. Cambridge University.

McNeish, D. M. (2015). Using Lasso for Predictor Selection and to Assuage Overfitting: A Method Long Overlooked in Behavioral Sciences. *Multivariate Behavioral Research*, *50*(5), 471-484.

Mikalef, P., & Pateli, A. (2017). Information Technology-enabled Dynamic Capabilities and Their Indirect Effect on Competitive Performance: Findings from PLS-SEM and fsQCA. *Journal of Business Research*, *70*, 1-16.

Mingers, J. (2001). Combining IS Research Methods: Towards a Pluralist Methodology. *Information Systems Research*, *12*(3), 240-259.

Mingers, J. (2003). The Paucity of Multimethod Research: A Review of the Information Systems Literature. *Information Systems Journal*, *13*(3), 233-249.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A Manifesto for Reproducible Science. *Nature Human Behaviour*, *1*(1), 0021.

Nagel, E. (1979). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Hackett.

Nicolai, A., & Seidl, D. (2010). That's Relevant! Different Forms of Practical Relevance in Management Science. *Organization Studies*, *31*(9-10), 1257-1285.

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The Preregistration Revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600-2606.

Nuzzo, R. (2014). Scientific Method: Statistical Errors. *Nature*, *506*(7487), 150-152.

Nuzzo, R. (2015). How Scientists Fool Themselves – And How They Can Stop. *Nature*, *526*, 182-185.

Open Science Collaboration. (2015). Estimating the Reproducibility of Psychological Science. *Science*, *349*(6251), aac4716.

Pagell, M. (2021). Replication Without Repeating Ourselves: Addressing the Replication Crisis in Operations and Supply Chain Management Research. *Journal of Operations Management*, *67*(1), 105-115.

Petter, S., & Hadavi, Y. (2021). With Great Power Comes Great Responsibility: The Use of Partial Least Squares in Information Systems Research. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, *52*(SI), 10-23.

Pitt, M. A., & Myung, I. J. (2002). When A Good Fit Can Be Bad. *Trends in Cognitive Sciences*, *6*(10), 421-425.

Platt, J. R. (1964). Strong Inference: Certain Systematic Methods of Scientific Thinking May Produce Much More Rapid Progress Than Others. *Science*, *146*(3642), 347-353.

Popper, K. R. (1959). *The Logic of Scientific Discovery*. Routledge.

Preacher, K. J., & Merkle, E. C. (2012). The Problem of Model Selection Uncertainty in Structural Equation Modeling. *Psychological Methods*, *17*(1), 1-14.

Prosperi, M., Bian, J., Buchan, I. E., Koopman, J. S., Sperrin, M., & Wang, M. (2019). Raiders of the Lost HARK: A Reproducible Inference Framework for Big Data Science. *Palgrave Communications*, *5*(1), 125.

Ragin, C. C. (2008). *User's Guide to Fuzzy-Set / Qualitative Comparative Analysis* University of Arizona.

Rasoolimanesh, S. M., Ringle, C. M., Sarstedt, M., & Olya, H. (2021). The Combined Use of Symmetric and Asymmetric Approaches: Partial Least Squares-structural Equation Modeling and Fuzzy-set Qualitative Comparative Analysis. *International Journal of Contemporary Hospitality Management*, *33*(5), 1571-1592.

Richter, N. F., Cepeda Carrión, G., Roldán, J. L., & Ringle, C. M. (2016a). European Management Research Using Partial Least Squares Structural Equation Modeling (PLS-SEM): Editorial. *European Management Journal*, *34*(6), 589-597.

Richter, N. F., Hauff, S., Ringle, C. M., & Gudergan, S. P. (2022). The Use of Partial Least Squares Structural Equation Modeling and Complementary Methods in International Management Research. *Management International Review*, *62*, 449-470.

Richter, N. F., Hauff, S., Ringle, C. M., Sarstedt, M., Kolev, A. E., & Schubring, S. (2023). How to Apply Necessary Condition Analysis in PLS-SEM. In H. Latan, J. J. F. Hair, & R. Noonan (Eds.), *Partial Least Squares Path Modeling: Basic Concepts, Methodological Issues and Applications* (pp. 267-297). Springer International Publishing.

Richter, N. F., Sinkovics, R. R., Ringle, C. M., & Schlägel, C. (2016b). A Critical Look at the Use of SEM in International Business Research *International Marketing Review*, *33*(3), 376-404.

Richter, N. F., & Tudoran, A. A. (2024). Elevating Theoretical Insight and Predictive Accuracy in Business Research: Combining PLS-SEM and Selected Machine Learning Algorithms. *Journal of Business Research*, *173*, 114453.

Rigdon, E., Sarstedt, M., & Moisescu, O. I. (2023). Quantifying Model Selection Uncertainty via Bootstrapping and Akaike Weights. *International Journal of Consumer Studies*, *47*(4), 1596-1608.

Rigdon, E. E. (2024). Understanding Composite-Based Structural Equation Modeling Methods From the Perspective of Regression Component Analysis. *Multivariate Behavioral Research*, *forthcoming*.

Rigdon, E. E., Sarstedt, M., & Ringle, C. M. (2017). On Comparing Results from CB-SEM and PLS-SEM. Five Perspectives and Five Recommendations. *Marketing ZFP*, *39*(3), 4-16.

Ringle, C. M., Sarstedt, M., & Schlittgen, R. (2014). Genetic Algorithm Segmentation in Partial Least Squares Structural Equation Modeling. *OR Spectrum*, *36*(1), 251-276.

Ringle, C. M., Sarstedt, M., Schlittgen, R., & Taylor, C. R. (2013). PLS Path Modeling and Evolutionary Segmentation. *Journal of Business Research*, *66*(9), 1318-1324.

Ringle, C. M., Sarstedt, M., Sinkovics, N., & Sinkovics, R. R. (2023). A Perspective on Using Partial Least Squares Structural Equation Modelling in Data Articles. *Data in Brief*, *48*, 109074.

Robey, D. (1996). Research Commentary: Diversity in Information Systems Research: Threat, Promise, and Responsibility. *Information Systems Research*, *7*(4), 400-408.

Rubin, M., & Donkin, C. (2022). Exploratory Hypothesis Tests Can Be More Compelling Than Confirmatory Hypothesis Tests. *Philosophical Psychology*, 1-29.

Sánchez, G. (2009). *PATHMOX Approach: Segmentation Trees in Partial Least Squares Path Modeling (Doctoral Dissertation)*. Universitat Politècnica de Catalunya.

Sarstedt, M., & Adler, S. J. (2023). An Advanced Method to Streamline P-hacking. *Journal of Business Research*, *163*, 113942.

Sarstedt, M., Adler, S. J., Ringle, C. M., Cho, G., Diamantopoulos, A., Hwang, H., & Liengaard, B. D. (2024). Same Model, Same Data, But Different Outcomes: Evaluating the Impact of Method Choices in Structural Equation Modeling. *Journal of Product Innovation Management*.

Sarstedt, M., Becker, J.-M., Ringle, C. M., & Schwaiger, M. (2011). Uncovering and Treating Unobserved Heterogeneity with FIMIX-PLS: Which Model Selection Criterion Provides an Appropriate Number of Segments? *Schmalenbach Business Review*, *63*(1), 34-62.

Sarstedt, M., & Danks, N. P. (2022). Prediction in HRM Research: A Gap Between Rhetoric and Reality. *Human Resource Management Journal*, *32*(2), 485-513.

Sarstedt, M., Hair, J. F., Pick, M., Liengaard, B. D., Radomir, L., & Ringle, C. M. (2022a). Progress in Partial Least Squares Structural Equation Modeling Use in Marketing Research in the Last Decade. *Psychology & Marketing*, *39*(5), 1035-1064.

Sarstedt, M., & Moisescu, O.-I. (2023). Quantifying Uncertainty in PLS-SEM-based Mediation Analyses. *Journal of Marketing Analytics*.

Sarstedt, M., Radomir, L., Moisescu, O. I., & Ringle, C. M. (2022b). Latent Class Analysis in PLS-SEM: A Review and Recommendations for Future Applications. *Journal of Business Research*, *138*, 398-407.

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2017). Treating Unobserved Heterogeneity in PLS-SEM: A Multi-Method Approach. In R. Noonan & H. Latan (Eds.), *Partial Least Squares Structural Equation Modeling: Basic Concepts, Methodological Issues and Applications* (pp. 197-217). Springer.

Sarstedt, M., Ringle, C. M., Henseler, J., & Hair, J. F. (2014). On the Emancipation of PLS-SEM: A Commentary on Rigdon (2012). *Long Range Planning*, *47*(3), 154-160.

Schlittgen, R., Ringle, C. M., Sarstedt, M., & Becker, J.-M. (2016). Segmentation of PLS Path Models by Iterative Reweighted Regressions. *Journal of Business Research*, *69*(10), 4583-4592.

Schuberth, F., Rademaker, M. E., & Henseler, J. (2023). Assessing the Overall Fit of Composite Models Estimated by Partial Least Squares Path Modeling. *European Journal of Marketing*, *57*(6), 1678-1702.

Schwab, S., & Held, L. (2020). Different Worlds: Confirmatory Versus Exploratory Research. *Significance*, *17*(2), 8-9.

Schwarz, G. (1978). Estimating the Dimensions of a Model. *Annals of Statistics*, *6*(2), 461-464.

Sharma, P. N., Liengaard, B. D., Hair, J. F., Sarstedt, M., & Ringle, C. M. (2023). Predictive Model Assessment and Selection in Composite-based Modeling Using PLS-SEM: Extensions and Guidelines for Using CVPAT. *European Journal of Marketing*, *57*(6), 1662-1677.

Sharma, P. N., Sarstedt, M., Shmueli, G., Kim, K. H., & Thiele, K. O. (2019). PLS-Based Model Selection: The Role of Alternative Explanations in Information Systems Research. *Journal of the Association for Information Systems*, *20*(4), 346-397.

Sharma, P. N., Shmueli, G., Sarstedt, M., Danks, N., & Ray, S. (2021). Prediction-oriented Model Selection in Partial Least Squares Path Modeling. *Decision Sciences*, *52*(3), 567-607.

Shiau, W.-L., Chau, P. Y. K., Thatcher, J. B., Teng, C.-I., & Dwivedi, Y. K. (2024). Have We Controlled Properly? Problems With and Recommendations for the Use of Control Variables in Information Systems Research. *International Journal of Information Management*, *74*, 102702.

Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, *25*(3), 289-310.

Shmueli, G., & Koppius, O. R. (2011). Predictive Analytics in Information Systems Research. *MIS Quarterly*, *35*(3), 553-572.

Shmueli, G., Ray, S., Velasquez Estrada, J. M., & Chatla, S. B. (2016). The Elephant in the Room: Evaluating the Predictive Performance of PLS Models. *Journal of Business Research*, *69*(10), 4552-4564.

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive Model Assessment in PLS-SEM: Guidelines for Using PLSpredict. *European Journal of Marketing*, *53*(11), 2322-2347.

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. *Annual Review of Psychology*, *69*, 487-510.

Stebbins, R. (2001). *Exploratory Research in the Social Sciences*. Sage.

Steyerberg, E. W., & Harrell, F. E. (2016). Prediction Models Need Appropriate Internal, Internal-external, and External validation. *Journal of Clinical Epidemiology*, *69*, 245-247.

Straub, D. (2012). Editor's Comments: Does MIS Have Native Theories? *MIS Quarterly*, *36*(2), iii-xii.

Straub, D., & Ang, S. (2011). Editor's Comments: Rigor and Relevance in IS Research: Redefining the Debate and a Call for Future Research. *MIS Quarterly*, *35*(1), iii-xi.

Sullivan, G. M., & Feinn, R. (2012). Using Effect Size-or Why the P Value Is Not Enough. *Journal of Graduate Medical Education*, *4*(3), 279-282.

Sutton, R. I., & Staw, B. M. (1995). What Theory is Not. *Administrative Science Quarterly*, *40*(3), 371-384.

Swedberg, R. (2020). Exploratory Research. In C. Elman, J. Gerring, & J. Mahoney (Eds.), *The Production of Knowledge: Enhancing Progress in Social Science* (pp. 17-41). Cambridge University Press.

Szollosi, A., Kellen, D., Navarro, D. J., Shiffrin, R., van Rooij, I., Van Zandt, T., & Donkin, C. (2020). Is Preregistration Worthwhile? *Trends in Cognitive Sciences*, *24*(2), 94-95.

Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y.-M., & Lauro, C. (2005). PLS Path Modeling. *Computational Statistics & Data Analysis*, *48*(1), 159-205.

Tomarken, A. J., & Waller, N. G. (2003). Potential Problems with "Well Fitting" Models. *Journal of Abnormal Psychology*, *112*(4), 578-598.

Tukey, J. W. (1980). We Need Both Exploratory and Confirmatory. *The American Statistician*, *34*(1), 23-25.

Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems. *MIS Quarterly*, *37*(1), 21-54.

Venkatesh, V., Brown, S. A., Maruping, L. M., & Bala, H. (2008). Predicting Different Conceptualizations of System Use: The Competing Roles of Behavioral Intention, Facilitating Conditions, and Behavioral Expectation. *MIS Quarterly*, *32*(3), 483-502.

Venkatesh, V., Brown, S. A., & Sullivan, Y. W. (2016). Guidelines for Conducting Mixed-methods Research: An Extension and Illustration. *Journal of the Association for Information Systems*, *17*(7), 435-494.

Wager, S., & Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, *113*(523), 1228-1242.

Ward, M. D., Greenhill, B. D., & Bakke, K. M. (2010). The Perils of Policy by p-value: Predicting Civil Conflicts. *Journal of Peace Research*, *47*(4), 363-375.

Weick, K. E. (1989). Theory Construction as Disciplined Imagination. *The Academy of Management Review*, *14*(4), 516-531.

Weick, K. E. (1995). What Theory is Not, Theorizing Is. *Administrative Science Quarterly*, *40*(3), 385-390.

Wold, H. (1982). Soft Modeling: The Basic Design and Some Extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems Under Indirect Observations: Part II* (pp. 1-54). North-Holland.

Yan, Y., Chen, H., Shao, B., & Lei, Y. (2023). How IT Affordances Influence Customer Engagement in Live Streaming Commerce? A Dual-stage Analysis of PLS-SEM and fsQCA. *Journal of Retailing and Consumer Services*, *74*, 103390.

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspectives on Psychological Science*, *12*(6), 1100-1122.