Guidelines for the Development of Three-Level Models: Bridging Levels of Analysis and Integrating Contextual Influences in IS Research

Viswanath Venkatesh, Qin Weng, Arun Rai, Likoebe M. Maruping

Abstract

The use of multilevel analysis has steadily increased in information systems (IS) research. Many studies are doing an admirable job of integrating two-level models into their examination of IS phenomena. However, two-level models are limited in how well they enable researchers to (1) more explicitly incorporate context into theory development and testing and (2) bridge the existing gap between micro- and macro-level research by focusing on intervening mechanisms that link hierarchically distal levels of analysis. Three-level models have emerged as a potential way to address these limitations. While literature has clearly outlined the mechanics of how to estimate three-level models, there is very little, if any, guidance on when and how to integrate the use of such models with theory development. Consequently, IS researchers have little guidance to inform their decisions about integrating the use of three-level models with their theory development and testing. In this article, we identify the circumstances under which IS researchers should consider the use of three-level models, develop guidelines about how to map the use of three-level model estimation to the theoretical objectives, and provide an illustration of how to implement the guidelines.

Keywords: three-level models, multilevel analysis, hierarchical structure, IS research, cross-level

1 Introduction

Many of the phenomena we examine in information systems (IS) research unfold at multiple levels of analysis (Tarasewich & Warkentin, 2002). For instance, employees who use an information system in an organization may be nested within teams or departments that have adopted a particular information system (Ahuja & Thatcher, 2005; Burton-Jones & Gallivan, 2007; Thatcher et al., 2006). Developers working on software projects may be nested within project teams (Colazo & Fang, 2010; Medappa & Srivastava, 2019; Rai et al., 2009). Similarly, from a macro-perspective, organizations may be nested within strategic groups which in turn may be nested within industries (Chiasson & Davidson, 2015) or countries (Srivastava et al., 2016; Srivastava & Teo, 2007, 2008, 2010). This nested structure of the entities (e.g., individuals, groups, subunits, organizations, industries, countries) that are the focus of much IS research creates the potential for examining relationships that cross levels of analysis such that factors associated with entities at a higher level of analysis (e.g., a business unit, industry) have an influence on the outcomes of entities nested within those higher-level units (e.g., employees within the business unit, firms within the industry). The opportunity to enrich our theories and understanding of IS phenomena by explicitly modeling this hierarchically nested structure of data has increased as the tools and methodologies for analyzing such data have become more accessible. This has enabled researchers to integrate micro- and macro-level theories to create meso-level theories (Bamberger, 2008; Bauer et al., 2006; House et al., 1995; Klein et al., 1994; Mathieu & Chen, 2011; Morgeson & Hofmann, 1999; Rousseau, 2011).

As a result of advancements in the tools for analyzing hierarchically nested data, the IS literature has seen the emergence of an increasing number of multilevel studies (Bélanger et al., 2014). Researchers have examined IS phenomena, such as technology implementation and use (Burton-Jones & Gallivan, 2007; Kang et al., 2012; Sasidharan et al., 2012), resistance to implementation (Kane & Labianca, 2011; Lapointe & Rivard, 2005), post-implementation and post-adoption use (Maruping & Magni, 2015), IS project control (Kirsch et al., 2010), project risk management (Windeler et al., 2017), offshore software development (Rai et al., 2009), IT employee compensation (Ang et al., 2002), and e-government (Srivastava et al., 2016). These studies have served as important stepping-stones to yield rich theoretical insights on IS phenomena that have traditionally been examined at a single level of analysis. Almost all of these studies have focused on examining phenomena at two levels of analysis. In recent years, there has been an increase in the number of three-level studies. Specifically, we reviewed papers published in 8 leading IS journals[1] between 2011 and 2020, and found 10 quantitative papers that

[1] These eight journals include MIS Quarterly, Information Systems Research, Journal of Management Information Systems, Journal of the AIS, European Journal of Information Systems, Information Systems Journal, Journal of Information Technology, and Journal of Strategic Information Systems. Further details of the three-level papers we identified in these journals are provided in Appendix C.

utilized three-level analysis (i.e., Brohman et al., 2020; Kane & Borgatti, 2011; Kane & Labianca, 2011; Ma et al., 2014; Sasidharan et al., 2012; Venkatesh et al., 2018, 2020; Wang et al., 2019; Xie & Lee, 2015; Zhao et al., 2018).

Although valuable, the examination of two-level models has some limitations in facilitating the bridging of the macro-micro divide in IS research. Two particularly important limitations are apparent. First, two-level models are limited in how they account for contextual influences in cross-level relationships (e.g., how team expertise integration affects individual learning across different departmental units). Further, as evidenced by our literature review of multilevel papers, there is an increasing need to capture changes over time, which by itself introduces an additional level of analysis (e.g., longitudinal changes at the individual level, nested within a group—which makes for time nested within individuals nested within teams). Seven out of the 10 three-level papers in our review are of this type (Appendix C). Two-level models are limited in their ability to incorporate such a contextual level/role for time-variant individual data nested within teams, whereas three-level models allow us to capture individual dynamics nested within a higher-level context. Second, two-level models are limited in their ability to account for intervening mechanisms that connect hierarchically distal units (e.g., how intermediate-level mechanisms link strategic business unit characteristics and individual employee system use). As we will discuss later in this article, three-level models have the ability to address some of these limitations. Three-level models extend two-level models by including an additional hierarchical level of nesting and, consequently, an ability to account for variation in lower-level outcomes that are attributable to factors at higher levels of the hierarchy (Raudenbush & Bryk, 2002; Raykov, 2010). We believe that, compared to two-level models, three-level models offer greater potential for IS researchers to pursue "blue ocean ideas" discussed by Straub (2009) across a wide spectrum of IS topics through the ability to more explicitly incorporate context into theory development and testing, and a greater ability to bridge factors that operate at macro-, meso-, and micro-levels of analysis.

The multilevel modeling literature has offered useful guidelines on how to analyze data that are nested across three different levels of analysis by explaining how to derive the system of equations (e.g., Raudenbush & Bryk, 2002), estimate the coefficients and variance components (e.g., Raykov, 2010; Yau et al., 1993), and conduct multilevel mediation and moderation (e.g., Pituch et al., 2010; Preacher, 2011; Raudenbush & Bryk, 2002). This literature has given researchers a useful basis for understanding the mechanics behind the estimation of three-level models. However, while research on how to conduct three-level multilevel analysis is quite robust, guidance on *when* and *how* to integrate its use with theory development and testing is lacking in comparison. Recently, Zhang and Gable (2017) offered guidelines on multilevel research. Although a key aid to researchers using multilevel models, the focus of their paper was on theorizing, rather than model development and application. Further, the examples provided in their guidelines were focused on two-level models. Hong et al. (2014) developed six-step guidelines for context-specific theorizing in IS research and illustrated the application through variations of the technology acceptance model (TAM; Venkatesh et al., 2003) in two technology use contexts (i.e., a digital library and an agile web portal). Although helpful for researchers to identify and incorporate context-specific factors in theorizing about IT artifacts, Hong et al. focused on single-context theory contextualization and involved little discussion on how to use multilevel modeling to incorporate contextual factors theoretically and methodologically. We find little, if any, research explaining the circumstances under which the use of three-level models is necessary or desirable in either accounting for contextual influences or bridging levels of analysis for IS phenomenon. Consequently, IS researchers have no guidance to inform their decision making about whether and how to integrate three-level models into theory development and testing. Considering this gap in the literature, the objectives of this research are to:

1) provide guidance about the circumstances in which IS researchers should consider integrating the use of three-level models in their theories,
2) develop guidelines on how to ensure adequate mapping of theory to three-level model specification via examination of cross-level main effects, cross-level moderation, and cross-level mediation testing,
3) provide an illustration of our proposed guidelines.

By accomplishing these objectives, we expect this research to make two important contributions. First, we go beyond the current multilevel modeling literature by integrating the three-level model estimation procedures with theory development and testing. This gives IS and other researchers tools that can inform their decision making about when and how to incorporate context and bridge different levels of analysis. Second, the guidelines enable researchers to develop richer, more comprehensive theories in IS and other fields that incorporate factors at higher, intervening, and lower levels of analysis and identify the mechanisms that link these different levels.

2 Background on Multilevel Research

Before delving into our discussion of three-level models, it is important to identify some common terminology and assumptions associated with multilevel research. The first has to do with the meaning of multilevel, the second pertains to the direction in which multilevel phenomena unfold, and the third relates to approaches to developing and testing multilevel models.

First, the term "multilevel" has several different meanings. The organization science literature has used the term multilevel synonymously with meso, multiple levels, and mixed levels, which incorporate notions of cross-level relationships, homology of relationships across levels of analysis and isomorphism of constructs across levels of analysis (Dansereau et al., 1984; House et al., 1995; Klein et al., 1994; Rousseau, 1985). More recently, multilevel research has been used more broadly to refer to research that incorporates different levels of analysis (Burton-Jones & Gallivan, 2007; Kozlowski & Klein, 2000; Mathieu & Chen, 2011). We adopt this recent view of multilevel research in this paper.

Second, the multilevel literature identifies two broad approaches to conceptualizing the direction in which cross-level relationships unfold: top-down and bottom-up (Kozlowski & Klein, 2000). In this paper, we focus on the most prevalent form examined in multilevel research in organization science—the top-down multilevel model; which allows researchers to incorporate contextual influences (Kozlowski & Klein, 2000). Broadly, higher-level units are theorized to influence lower-level units in a variety of ways (Kozlowski & Klein, 2000; Mathieu & Chen, 2011). Bottom-up or emergence multilevel models are beyond the scope of this paper.

Finally, the multilevel literature has used different approaches to developing and testing multilevel models. In this paper, we focus on random coefficient model (RCM) techniques. As noted by Burton-Jones and Gallivan, multilevel research has traditionally employed assumptions embedded in the "functionalist, positivist, and variance-oriented" (2007, p. 659) view of theory. As noted earlier, other approaches to multilevel theory development and testing have been utilized in IS research with great success (e.g., Lapointe & Rivard, 2005; Nan, 2011). However, RCM techniques that enable the development and testing of three-level models are embedded within the functionalist, positivist, and variance-oriented view. Therefore, consistent with much of the multilevel IS research (e.g., Ang et al., 2002; Boh et al., 2007; Burton-Jones & Gallivan, 2007; Rai et al., 2009) we adopt this approach while noting that there are other approaches that are equally valid and useful for pursuing multilevel theory development and testing.

3 Scientific Value of Three-Level Models

Three-level modeling is especially relevant for IS research, as the core thesis of the IS discipline revolves around the use of technology. As stated by Burton-Jones and Gallivan (2007), research that studies system use at one level at a time may suffer from a level bias, leading to "an unnatural, incomplete, and very disjointed view of how information systems are used in practice" (p. 657). Accordingly, they suggested that system use at any level of analysis comprises three elements including user, system, and task. Three-level modeling makes it possible for IS researchers to effectively address the complexities of the relationships between units at the levels these elements reside. Hong et al. (2014) further highlighted the importance of multilevel modeling in capturing the contexts where technologies are studied and used. They pointed out that technologies are always used in a specific context, and the use experiences may differ by users and use contexts. Three-level modeling provides a means to bring richer context to IS research as it can incorporate important constructs from different levels, and thus more richly and accurately represent the interplay between technologies, the users, and the use context. Together, as succinctly stated by Burton-Jones and Gallivan (2007), multilevel modeling "opens new opportunities for theory" and "may even help generate new organization-specific rather than reference-discipline-specific theories" (p. 660).

As noted in the introduction, we believe that the value of three-level models in multilevel theory development and testing lies in three broad roles: 1) developing richer theories about context in IS research by including theoretical perspectives and associated factors that operate across three levels of analysis, 2) bridging factors across three levels of analysis to more fully understand IS phenomena, and 3) integrating a mix of mechanisms (i.e., cross-level main, mediation, and moderation effects) that operate across three levels of analysis to model the complexity of hierarchically nested social systems. We discuss each of these in greater detail next, in an effort to offer IS researchers with guidance on why they might consider developing and testing three-level models.

3.1 Multilevel Theorizing and the Role of Context

The role of context in theory development and testing has evolved over time (Johns, 2001, 2006). As Figure 1 illustrates, at one end of the spectrum (shown in the two-level model), context emerges as a factor to be empirically considered with regard to generalizability and external validity of study results. In this role, context represents the setting in which the researcher collects data or observes a phenomenon unfold and is typically described in great depth in the method section of a paper. This approach would be consistent with what Johns (2006) referred to as "context as a constant," where context does not play an important role in the development and testing of a theory. As in other disciplines, this approach to context is highly prevalent in IS research where the research design or the use of control variables are used to control for differences in context. At the other end of the spectrum (shown in the three-level model), context emerges as a theoretically important variable that informs theory development and understanding of a phenomenon (Cappelli & Sherer, 1991; Johns, 2001; Hong et al., 2014; Mowday & Sutton, 1993; Rousseau & Fried, 2001). In multilevel theory development and testing, this latter approach to incorporating context allows researchers to explicitly identify aspects of the higher-level context that affect lower-level relationships or outcomes (Griffin, 2007; Johns, 2006). Such approaches to incorporating context have been heralded for their ability to shed light on differences in relationship strengths and direction at a lower-level of analysis (Burton-Jones & Gallivan, 2007; Griffin, 2007; Johns, 2006). Such an approach to theorizing and testing context, that spans levels of analysis, is less common in IS research but is sorely needed. Indeed, emerging perspectives on information technology (IT) suggest that it is embedded in entities that span multiple levels of analysis—e.g., products, processes, organizations, economies, and countries (Agarwal & Lucas, 2005; Kohli & Grover, 2008; Srivastava et al., 2016; Tarasewich & Warkentin, 2002). Thus, it behooves IS researchers to incorporate context more effectively at the various levels of analysis to understand where and how the effects of IT are manifested.

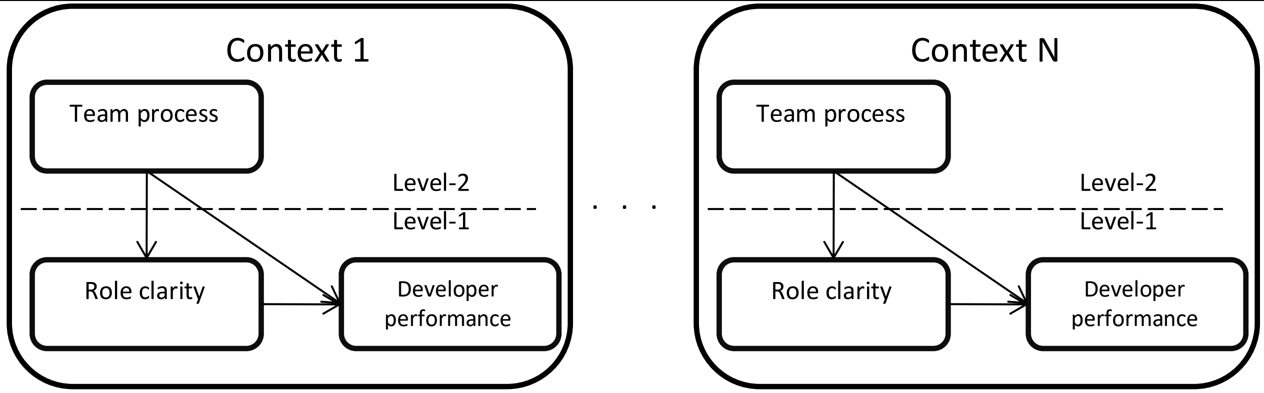
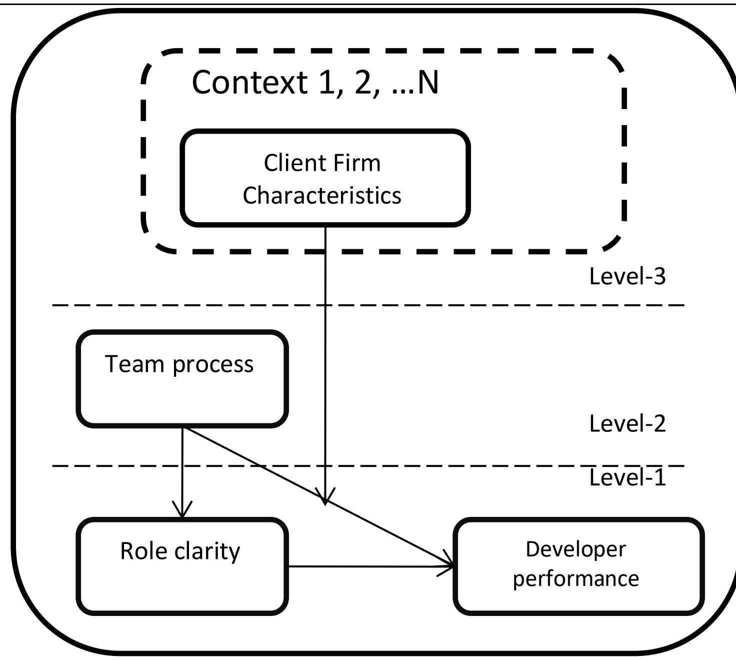
<table>
<tr><th>Model</th><th>Objective</th><th>Illustration</th></tr>
<tr><td>Two-level</td><td>Researcher incorporates hierarchically proximal level of context (team context) but not hierarchically distal level of context (client firm context)</td><td>
</td></tr>
<tr><td>Three-level</td><td>Researcher incorporates the effects of hierarchically proximal (team context) and distal context (client firm context)</td><td>
</td></tr>
</table>

Figure 1. Three-Level versus Two-Level Approaches to Incorporating Context into Multilevel Empirical Models

An obvious advantage that three-level models have, compared to two-level models, is that they give researchers the opportunity to more closely model the nested structure within which organizational and other social phenomena occur (Kozlowski & Klein, 2000; Misangyi et al., 2006; Roberts et al., 1978). That is, while two-level models also enable researchers to explicitly incorporate variations in context, they do so in a simplistic manner that may be limited in reflecting additional layers of context that are relevant to the phenomenon of interest. In particular, the ubiquity of IT across various levels of the organizational and societal hierarchy suggests that the IS phenomena that are of interest to IS researchers are very likely influenced by contextual elements at hierarchically proximal and distal levels of analysis. Consider a study on technology effectiveness in remote teaching. After the onset of COVID-19, many countries went through several rounds of lockdowns and schools moved classes to online instruction. Suppose the study examined the use of desktop educational software with students from schools in different countries as the focal phenomenon. In this case, there would be a significant omission if the study did not consider country-level variability because access to computers may be dramatically different across countries. For example, in some developing countries, although schools have computers, a typical household may not have a personal computer, thus constraining the effectiveness of the software. Access to the Internet, complementary platforms, and technologies may also vary across countries, which all potentially affect the use and effectiveness of the desktop educational software in different countries.

From a multilevel perspective, this presents an interesting case. Overall, we have three levels of variables to describe the phenomenon and the context in which it manifests: individual (level-1), school (level-2), and country (level-3). At the individual level, one might predict that teaching software quality (modeled at the first level) would predict intention to use the software.

Further, one might expect school incentive (modeled at the second level) to enhance the likelihood that software quality translates into actual learning outcomes (i.e., moderate the relationship between software quality and learning outcomes). Thus, our theoretical model includes a dependent variable software use intention (level-1), predicted by an independent variable perceived software quality (level-1), moderated by the variable school incentive (level-2). Although this may very well be true among consumers in developed countries, a researcher might find that school incentive plays no moderating role for students in developing countries, as many students may not even have access to computers at home. Failure to account for country context might lead to erroneous conclusions about the moderating role of school incentive—i.e., observing spurious relationships between variables at lower levels due to their failure to empirically account for higher-level contextual influences that affect the relationship (Burton-Jones & Gallivan, 2007; Hackman, 2003; Rousseau, 1985). By taking country context into account the researcher is able to account for variations in the moderating role of school incentive. A researcher could also theorize more richly about country context by including a context-specific variable such as a country's telecommunications infrastructure (modeled at the third level) to explain variation in the moderating role of school incentive. Such *theoretically-relevant* elements constitute what Johns refers to as discrete context—"situational variables that influence behavior directly or moderate relationships between variables" (2006, p. 393). Such approaches to incorporating higher-level context go a long way in averting the internal validity threat associated with committing a contextual fallacy (Rousseau, 1985). Two-level models would be limited in their ability to adequately model this reality of hierarchical nesting within a context.

If location-specific factors account for variability in lower-level outcomes, or even the cross-level relationship between business unit-level factors and lower-level outcomes, it is clear that a two-level model would require some simplifying assumptions that may not be reflective of the real-world context in which the phenomenon of interest is occurring. Hackman (2003) warned against the dangers of such simplifying assumptions when phenomena occur within a hierarchically nested context. Thus, similar to single-level models, two-level models carry the potential risk of contextual fallacies, which constitute a serious threat to internal validity (Rousseau, 1985). In this regard, Johns underscored the vital role that cross-level models play in advancing the theoretical role of context, noting that cross-level research is likely to "elucidate context when the discrete contextual levers that are thought to be responsible for context effects are explicitly theorized and measured" (2006, p. 401). Chiasson and Davidson (2015) expressed a similar sentiment in calling for more explicit theorizing on the role of industry context in IS research. Consistent with this objective, three-level models enable researchers to explicitly model the influence of contextual factors and yield richer theoretical insights about the mechanisms through which omnibus context affects phenomena of interest at lower levels of analysis (Johns, 2006).

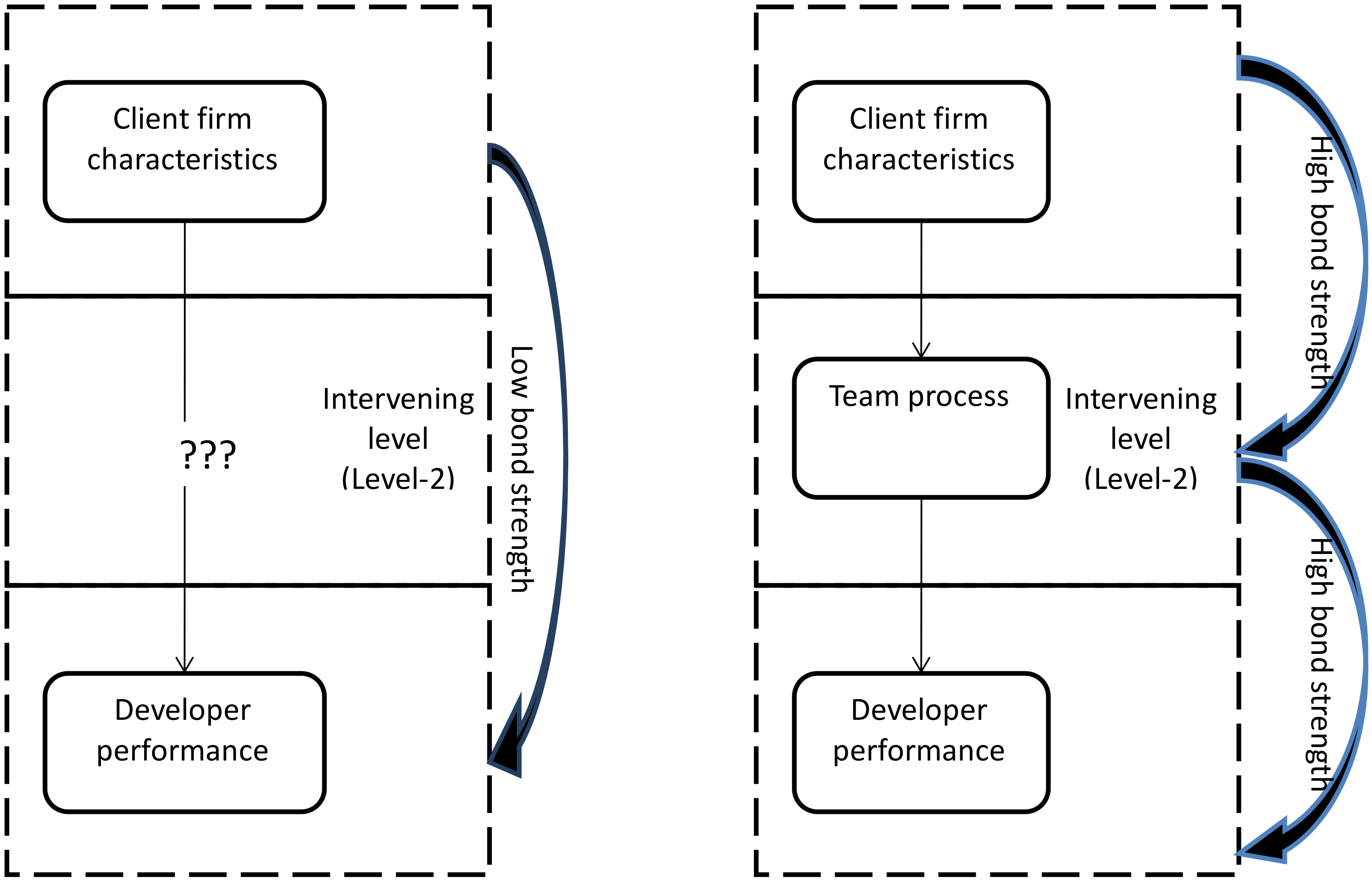
3.2 Multilevel Theorizing and Bridging Levels

In addition to providing the means to more explicitly incorporate context into multilevel theory development and testing, three-level models have the potential to play a key role in facilitating efforts to develop and test theory that bridges levels of analysis. As noted earlier, recent research has made significant strides in advancing theory using two-level models to test cross-level relationships (Zhang & Gable, 2017). However, there are a few limitations associated with two-level models that can be overcome through three-level models. We discuss each of these limitations in turn in the hopes that IS researchers will be better informed about when they might consider three-level versus two-level models.

First, while two-level models enable researchers to draw and test links between higher-level predictors and lower-level outcomes, there are theoretical limits to how this is accomplished. In an effort to incorporate a multiplicity of theoretical levels of analysis in multilevel theory development and testing, it is often tempting to connect hierarchically distal macro- and micro-level factors (Rousseau, 2011). For instance, in a study involving virtual teams nested within organizations, IS researchers might be interested in understanding how organizational structure affects the relationship between virtual teams' use of information and communication technology (ICT) and virtual team performance. Statistically, two-level RCM would allow researchers to test such a relationship. However, such an approach to examining relationships between these hierarchically distal factors raises some concern (Hackman, 2003; Whetten, 2002). From a theoretical standpoint there are numerous layers of mechanisms (e.g., leadership style, team training, uniformity of resource deployment, departmental policy and procedures) that lie between the macro-level, where organizational structure is conceptualized, and the micro- or meso-level where virtual team ICT use and performance are conceptualized. Indeed, Hackman (2003) warned that researchers need to be careful to ensure not to skip intervening levels of analysis when developing cross-level linkages between constructs of interest, as such an approach risks theoretically proximal explanatory dynamics being overlooked. Specifically, the linkage among different levels exhibits different strength, also known as *bond strength* (Kozlowski & Klein, 2000), referring to the extent to which one level's characteristics influence those at another level. Other researchers use similar concepts, such as "coupling" (Weick, 1976), "inclusion" (House et al., 1995), or "embeddedness" (Kozlowski & Salas, 1997), to describe the linkage strength between levels. Skipping intervening levels has the potential to violate the principle of bond strength, where antecedents within levels, or at a more proximal level, exert more powerful influence on outcomes of interest than hierarchically distal antecedents (Kozlowski & Klein, 2000). Ignoring bond strength across levels of analysis runs the risk of theoretically important mechanisms that link hierarchically distal levels being overlooked. In analytical terms, Moerbeek (2004) showed that omitting intervening levels of analysis can lead to overestimation of standard errors, increasing the risk of Type II error. Omitting intervening levels of analysis can also result in underestimated standard errors, by failing to account for variance attributable to non-independence at intervening levels, resulting in Type I error (Bliese & Hanges, 2004; Heck & Thomas, 2020). Importantly, Hackman noted that skipping intervening levels of analysis when drawing links between such hierarchically distal factors is tantamount to "replacing explanation with speculation" (2003, p. 918). As illustrated in Figure 2, the intervening level (i.e., team at level 2), has high bond strength to both higher-level and lower-level variables. Even though the intervening level may not be of focal interest, three-level models give researchers the ability to overcome this limitation of potential low bond strength in two-level models by incorporating the influence of factors at these intervening levels of analysis in a manner that is comparatively more theoretically sound. A study by Misangyi et al. (2006), which incorporates business segment-, corporate-, and industry-level effects, serves as a good exemplar of using three-level models to bridge various levels of analysis.

Second, for researchers examining relationships at a single level of analysis, multilevel modeling offers the potential to probe and fully understand observed relationships. Such approaches are especially beneficial when seemingly counterintuitive or unexpected results emerge from one's analysis. Hackman (2003) advocated a multilevel approach which he referred to as "bracketing." Here, the researcher probes the context within which the observed relationships occur by incorporating considerations one level up from the focal level of analysis and then also probes internal dynamics one level below the focal level of analysis. For instance, if one were examining the relationship between group-level technology use and group performance and found counterintuitive, unexpected, or non-significant relationships, the bracketing approach would suggest that the researcher examine the context (one level up) within which the teams in the sample are embedded (e.g., business unit, external leader) and also examine factors pertaining to the individuals who make up the teams (one level down) so as to gain better perspective. Two-level models give researchers the means to incorporate one level above or one level below the focal level of analysis, but not both simultaneously. In contrast, three-level models enable researchers to incorporate considerations from the three levels simultaneously, which can be highly beneficial for gaining a holistic picture of a complex phenomenon (Hackman, 2003; Rousseau & Fried, 2001).

In sum, three-level models offer several advantages over two-level models. These advantages are both theoretical and analytical in nature. Understanding when the development and use of three-level models is warranted is useful for building sound multilevel theory. However, although there are plenty of resources on how to specify and interpret three-level models, we recognize that there is a gap in our understanding on connecting multilevel theory to the use of such models.



Step 1: Check whether the constructs involved in the theoretical model are located at different levels.
Step 2: Check whether there are intermediate levels between the two levels of interest that should be included.
Step 3: Two-level models with skipped intermediate levels have low bond strength between levels (as shown on the left). Adding the intermediate level (team process) resolves the low bond strength issue, as the intermediate level has high bond strength with both the higher- and lower-level variables, functioning as a glue between the two.

Figure 2. Bridging Hierarchically Distal Levels of Analysis Using Three-Level Models

3.3 Benefits of Three-Level Models Compared to OLS Models

We next briefly explain the advantages of three-level models compared to ordinary least squares (OLS) regressions and conducting OLS regressions for subgroups. An important assumption of OLS regression is the independence of the sample observations that is often violated with nested data because the sample observations from the same subgroup tend to be correlated. Consider surveillance technology adoption in a multinational company that has divisions located in several countries. In this example, we want to answer the question, how does surveillance at work affect employee satisfaction with the work environment? We therefore model a relationship using two levels of data (i.e., employee at level-1 and division at level-2). The dependent variable is employee satisfaction with the work environment (level-1). The independent variables include division-level (level-2) variables (i.e., surveillance camera brand, quantity, and installed locations) and employee-level (level-1) variables (i.e., awareness of the surveillance, privacy knowledge, and organizational commitment). Using an OLS regression at the individual level overlooks the fact that other variability may exist among the divisions (level-2) and countries (level-3), wherein privacy is valued differently. Without considering the heterogeneity of the data across divisional offices, an OLS regression is likely to generate misleading results.

One option in this example is to conduct OLS regressions for subgroups. Indeed, if the research question only concerns the relationship in a specific subgroup, e.g., technology adoption in U.S. divisional offices, one can conduct OLS regressions at the division level by aggregating individual-level variables. But interpreting the results calls for caution because generalization and parsimony may be compromised, which are important considerations in theory development (Gregor, 2006). A multilevel model can partition the variability into different levels and thus more accurately accounts for lower-level relationships. It further allows different error structures at different levels (Raudenbush & Bryk, 2002). Moreover, analyzing the subgroups of a dataset can only potentially handle a top-level categorical grouping variable, whereas multilevel modeling allows including different types of contextual variables, both categorical and continuous, at the higher level, as shown in the three-level papers in our literature review (Appendix C). For example, Venkatesh et al. (2020), in their study on ICT4D use in rural India, the level-3 variables, lead user centralities for strong tie and weak tie, are continuous.

In fact, the aforementioned advantages of three-level models over OLS regressions apply to all multilevel models, including models with four or even more levels. Multilevel models are essential to study today's global business, and to incorporate contexts and temporal dynamics in our research. We focus on three-level models as they serve as a stepping stone for researchers to examine complex IS phenomena that involve constructs at multiple levels. Although three-level models are built on two-level models, they greatly extend the capabilities of two-level models by incorporating higher-level contexts, including when the study focuses on more dynamic and longitudinal data. Further, when two-level models include two levels that are distal in nature (i.e., with low bond strength), three-level models can bridge these distal levels by including intervening levels and account for intermediate mechanisms. In the sections that follow, we outline several guidelines for drawing such connections and demonstrate the use of these guidelines with an empirical illustration.

4 Three-Level Models: Theory Development and Analysis

4.1 Guideline 1: Determining When Three-Level Models Are Appropriate

For the IS researcher who is interested in testing a two-level model, there are circumstances where it may be necessary to determine if contextual effects at the third level need to be considered. We underscore that this determination is primarily driven by ***theoretical*** considerations. The researcher must carefully consider the extent to which the phenomenon of interest is likely to be affected by a hierarchically nested structure (Klein et al., 1994; Kozlowski & Klein, 2000; Rousseau & Fried, 2001). Often, a rich description of the context can give some insight on how much multiple levels may matter in examining a phenomenon of interest (Hackman, 2003; Johns, 2001, 2006; Rousseau & Fried, 2001). Many, though not all, of the phenomena IS researchers examine occur within a hierarchically nested social system and three-level models provide the means with which to model the complexities of such phenomena. Theoretical consideration must be given to such issues as the extent to which there is homogeneity versus heterogeneity in constructs of interest within higher-level units and the degree to which one would expect there to be independence of lower-level units from higher-level units (Klein et al., 1994). For instance, if the researcher's theory suggests that software project teams are independent entities, this may be reflected in the fact that the value of a construct for one team is independent of the value of the same construct for teams in the same business unit (Klein et al., 1994). If this assumption is challenged, i.e., the team-level variable is non-independent, multilevel modeling should be considered to ensure the heterogeneity across teams is taken into account. Therefore, theoretical considerations serve as the primary factor in determining whether multilevel models, including three-level models, are appropriate.

We suggest researchers consider the possible nested structure of a phenomenon under investigation, which may potentially help with formulating interesting research questions. As pointed out by Alvesson and Sandberg (2011), a good way of problematizing is through identifying and challenging assumptions of existing theories. Rather than focusing on filling research gaps, it is important to articulate why "it is important to fill this gap" (Alvesson &

Sandberg, 2011, p. 250). Considering how assumptions at a lower level of theory and analysis can be challenged by considering higher levels that describe the heterogeneity in the context can be an effective way to formulate questions. Such assumption challenging can pertain to the meaning of constructs, the relationships among constructs and the effects on outcomes, and the mechanisms.

Staying mindful of multilevel structure can aid with "prescience" theorizing, a concept proposed by Corley and Gioia (2011) for emergent and future phenomena that are likely to be characterized by changing organizations and societal phenomenon. Examining multilevel factors in such scenarios creates opportunities to discover unknown relationships that can emerge in a new future state of a technology, organization, or social system. In sum, we suggest that the decision to use three-level modeling be grounded in theoretical considerations. Specifically, if the constructs essential to the theoretical model are situated in a nested structure, multilevel modeling should be explored and evaluated. Figure 3 details theoretical considerations to be taken when we have a two-level model which can potentially be expanded to a three-level model.

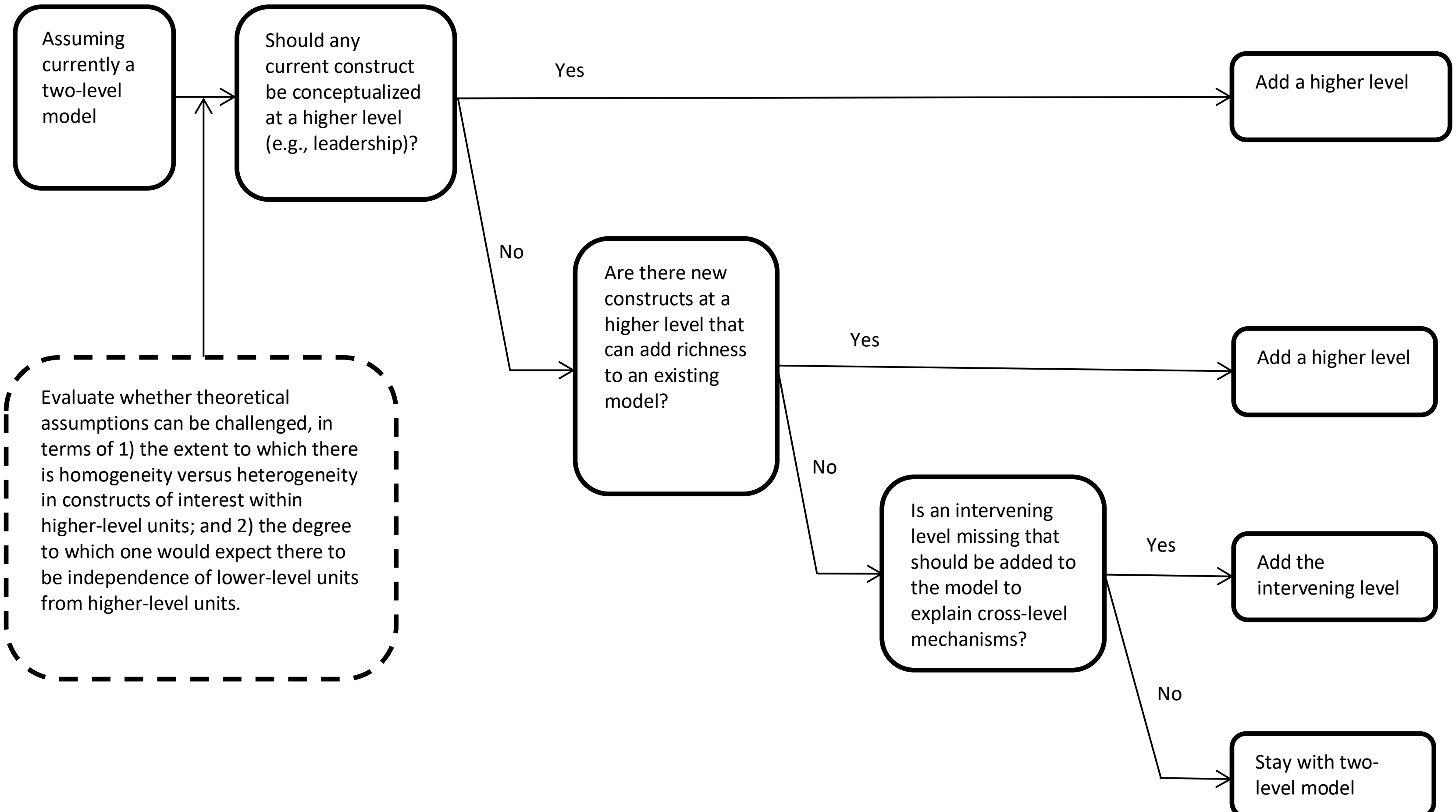


Figure 3. Theoretical Considerations When Deciding Whether to Expand a Two-Level Model to a Three-Level Model

The theoretical grounds on which the need for a three-level model is based, naturally translates to the structure of the data collected by the researcher. For instance, a researcher who collects data on 500 software developers, nested within 40 software project teams serving a single client firm of an organization may not need to incorporate a third level of analysis in examining relationships between project team processes (a level-2 factor) and individual developer performance (a level-1 outcome). This is simply because contextual factors at a higher level of analysis are held constant. However, if the IS researcher collects data on 500 software developers, nested within 40 software project teams across 15 client firms of an organization, it is clear that there may be theoretical grounds for considering a third level of analysis when testing the two-level model (e.g., if the constructs pertaining to developers and project teams can theoretically be expected to be homogeneous within client firms and heterogeneous between client firms). The researcher would need to consider the likelihood that elements at the client firm level (level-3) could be reasonably expected to affect the relationships of interest at the lower levels. In the context of this hypothetical study, perhaps client firm characteristics, such as country and size (level-3), may affect the cross-level relationship between software project team process and developer performance. Although the researcher may not have directly measured the actual characteristics of client firms, it would be important to determine if between-client firm differences account for variability in the lower-level outcomes. It could be that client firm characteristics are standardized across client firms such that there is little to no variability between them. However, it is also possible that there may be variability across client firms if they have different cultures and/or other characteristics. In short, it is important for the researcher to determine where the variability lies. This can be accomplished by decomposing the variance in the outcome of interest, developer performance, into its component parts (i.e., determining how much variability in the outcome resides at each level of analysis). The process for assessing variation across levels of analysis is discussed in greater detail in our illustration and specific equations are shown in Appendix A.

Once the researcher is able to determine that between-client firm differences account for variance in the outcome of interest, she can proceed to test the two-level model of interest while controlling for membership at the third level. As a further step, the researcher can determine whether treating project team-level coefficients as fixed effects versus random effects yields better model fit by conducting a likelihood-ratio test (Davison et al., 2002). This involves a simple comparison of deviance statistics from a restricted (fixed level-2 slopes and intercepts) and a less restricted (random level-2 slopes and intercepts) model (Raudenbush & Bryk, 2002). An unrestricted model where level-2 coefficients are found to vary across client firms would suggest that the effects of project team processes on developer-level outcomes are different across client firms. Compared to a simple two-level model (which assumes that cross-level effects of project team processes on developer performance are independent of the client firm to which the project team is assigned to), such an approach will ensure that unbiased estimates are obtained when testing the model of interest (Bliese & Hanges, 2004; Raykov, 2010). Failure to account for this third level of nesting can result in inflated standard errors and increase the risk of Type II error occurring. Figure 4 provides a summary of the factors to be considered when examining two-level models within a higher-level context.

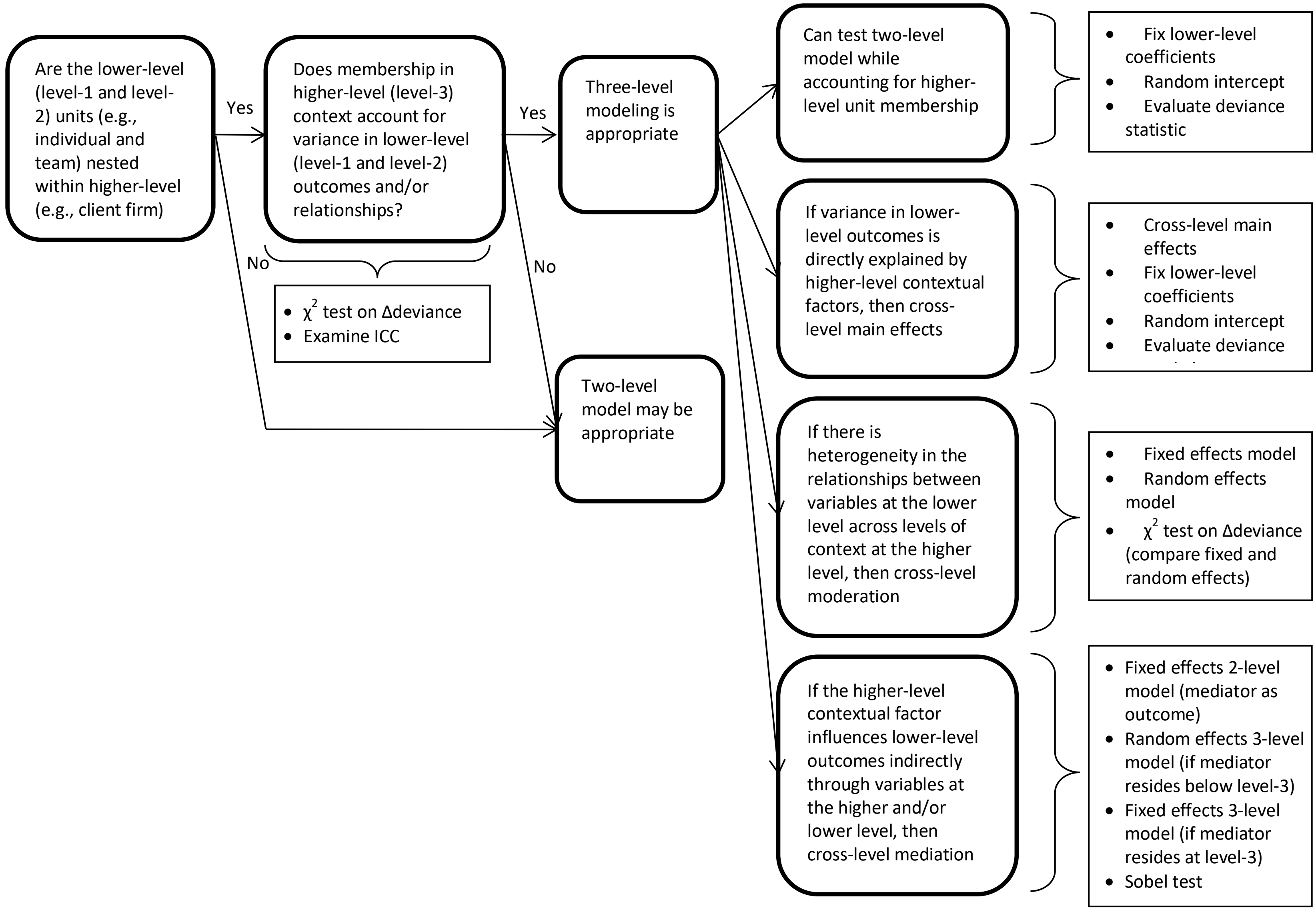


Figure 4. Methodological Considerations When Determining Whether Three-Level Modeling is Necessary

4.2 Guideline 2: Incorporating Factors at the Third Level

Once Guideline 1 has been completed, and the researcher has determined that variation in the outcome of interest exists across the three levels of analysis, the researcher can proceed to test the theoretical model of interest. Note that there are numerous variations on the types of relationships that might be examined here, including cross-level main effects, cross-level moderation, and cross-level mediation. Each of these types of relationship can take a variety of forms. For example, cross-level moderation could involve the moderating effect of a level-3 predictor of the relationship between a level-1 predictor and level-1 outcome or it could involve the moderating effect of a level-3 predictor between a level-2 predictor and a level-1 outcome, or even the interactive effect of two level-3 predictors on a level-1 outcome. Yet another variant on cross-level moderation could involve mediated moderation wherein the moderating effect of a level-3 factor on the relationship between two level-1 factors is mediated by an intervening factor at level-2. Similarly, cross-level mediation can take a variety of forms as will be discussed later. In the interest of clarity of exposition, we outline broad guidelines for cross-level main effects, simple forms of cross-level moderation and cross-level mediation next.

4.2.1 Guideline 2.1: Testing for Three-Level Main Effects Model

As the first step, we run a model with no predictor variables at any level, a model also known as fully unconditional model, in which we include only the mean value and error term at each of the three levels. The model informs us about the how variance is allocated across the three levels. After the variability is examined, we can introduce predictors at their respective levels, which ultimately predict the outcome at the lowest level (i.e., the developer level). For example, if our goal is to explain the psychological stress of developers from different project teams in different client firms, we can include developer demographics, personality and experience as predictors at the developer level (level-1), project risk and team cohesion as predictors at the team level (level-2), and country, firm size and diversity as predictors at the client firm level (level-3). Following Guideline 1, we can run a three-level model where lower-level slopes are fixed, and then run a three-level model where slopes of interest are random. Finally, we run a likelihood-ratio test to see whether including random slopes improves model fit. Of course, variables in the same level may interact with each other, most commonly seen in the lowest level. For example, in their study on online gambling behavior, Ma et al. (2014) included the interaction of website use behavior (regular and extended use) and money stakes at the individual observation level (level-3). Table 1 below illustrates the steps three-level main effects model.

Table 1. Summary of Guidelines for Three-Level Main Effects Models

Guideline	Objective	Steps
Guideline 1 (Variance components)	Determine where variability in lower-level (developer) outcome resides	• Step 1: Use a three-level unconditional model to decompose variance into three-level components (check whether the random effect is significant & whether ICC is higher than the threshold)
Guideline 2.1 (Main effects)	Model comparison and selection	• Step 1: Run a three-level unconditional model with control variables only • Step 2: Include main effects • Step 3: Include interaction terms (if any) at a specific level • Step 4: Compare model fit using deviance statistics

4.2.2 Guideline 2.2: Testing for Level-3 Cross-Level Moderation

As with two-level models, moderation in three-level models can be examined in two ways. The first form of cross-level moderation comes in the form of two factors at level-3 interacting to influence a level-1 outcome (Figure 5a). For instance, client firm size (level-3) might be hypothesized to interact with client firm decision-making autonomy (also level-3) in predicting developer performance (level-1). Here the interaction effect can be modeled consistent with the approach for such cross-level moderation in two-level models while accounting for project team membership at level-2. The second form of cross-level moderation occurs when a factor at level-3 moderates the relationship between variables at lower levels (level-1 or level-2) (Figures 5b and 5c). For instance, client firm characteristics might be hypothesized to moderate the relationship between project team processes and developer performance. Testing for the latter form of cross-level moderation involves two major steps. First, as suggested in Guideline 1 above, the researcher would need to establish that the slope of the lower-level coefficient varies across level-3 units. In this example, the researcher would want to demonstrate that the coefficient for project team processes varies across client firms. This can be accomplished by simply modeling the coefficient for project team processes as a random (rather than fixed) effect (Luke, 2004; Raudenbush & Bryk, 2002). A likelihood-ratio test can be conducted using the deviance statistics from the restricted and unrestricted model to determine if accounting for variation in level-2 slopes improves model fit (Davison et al., 2002; Raudenbush & Bryk, 2002). Once this is established to be the case, the researcher can then introduce level-3 predictors into the model that would explain variation in the slope of project team processes across client firms.

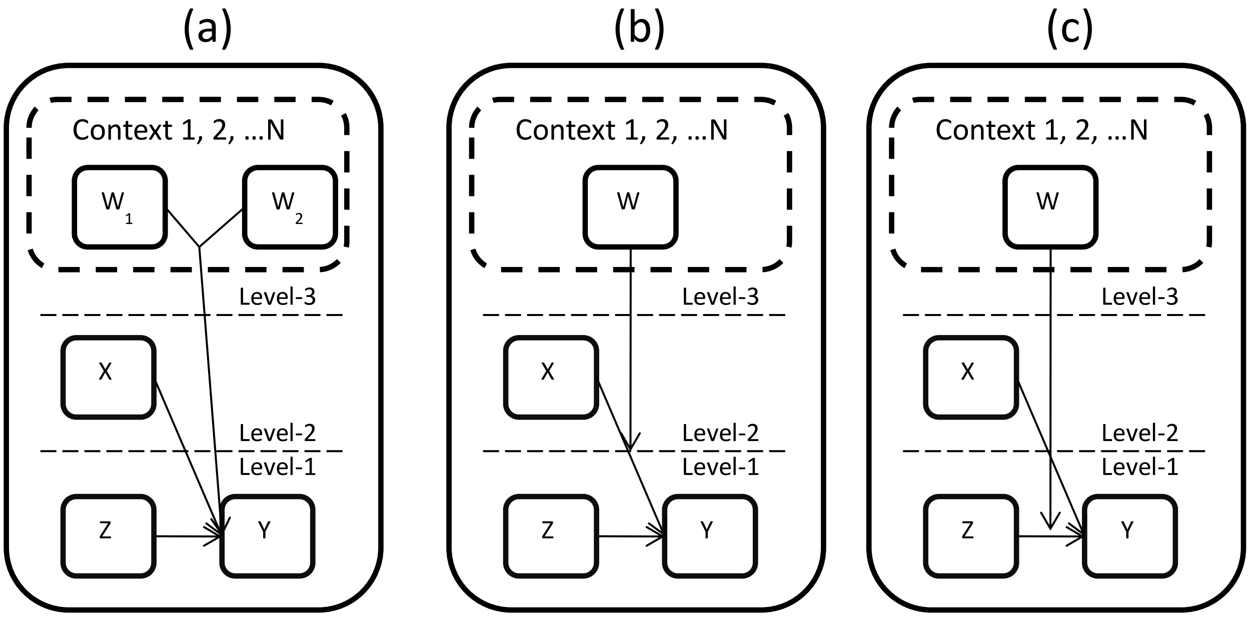


Figure 5. Different Forms of Cross-Level Moderation in Three-Level Models

As noted earlier, a similar approach can be used to determine if the relationship between variables at level-1 depends on a variable at level-3. For instance, the researcher might predict that the effect of a developer's mental model on task performance is contingent on the client firm culture. In this case, the researcher would determine if the slope for developer mental model varies across client firms while accounting for developer nesting within project teams at level-2. If significant variation is found in level-1 slopes across client firms, then level-3 predictors can be entered into the model. Similar to two-level models, the decrease in the deviance statistic can be examined to evaluate the extent to which the cross-level moderation improves model fit. A summary of the steps involved in testing the role of discrete context in three-level models is included in Table 2 below. We provide an example of how to test for cross-level moderation in three-level models in our illustration. More examples of cross-level moderation can also be found in four of the three-level papers identified in our literature review (Appendix C). For example, one of these studies examines how project management and process control affect individual outcomes (Venkatesh et al., 2018). In this study, level-2 variables (i.e., internal process control and external process control) moderate the relationships between level-2 predictors (requirement risk and project complexity risk) and level-1 outcomes (individual performance and psychological stress).

Table 2. Summary of Guidelines for Cross-Level Moderation in Three-Level Models

Guideline	Objective	Steps
Guideline 1 (Variance components)	Determine where variability in lower-level (developer) outcome resides	• Step 1: Use a three-level unconditional model to decompose variance into three-level components
Guideline 2.2 (Cross-level moderation)	Determine if lower-level relationships vary as a function of higher-level unit membership (client firm)	• Step 1: Follow Guideline 1 to determine where variance in outcome variable resides • Step 2: Run a three-level model where lower-level slopes are fixed • Step 3: Run a three-level model where lower-level slopes of interest are random • Step 4: Run a likelihood-ratio test to compare nested models for improvement in model fit
	Determine if lower-level relationships vary as a function of antecedents at higher-level unit (client firm level)	• Step 1: Include level-3 moderator • Step 2: Examine deviance statistics for improvement in model fit

4.2.3 Guideline 2.3: Testing for Cross-Level Mediation

Cross-level mediation in two-level models has received attention in the literature (Bauer et al., 2006; Chen et al., 2007; Mathieu & Taylor, 2007; Seibert et al., 2004; Zhang et al., 2009). Such cross-level models can take different forms. For instance, with upper-level mediation (also referred to as 2-2-1 mediation), the influence of a higher-level predictor on a lower-level outcome can be mediated (partially or fully) by an intervening variable at the higher level of analysis—e.g., the effect of project team autonomy (team-level) on developer performance (individual-level) being mediated by development methodology use (team-level). Alternatively, with lower-level mediation (also referred to as 2-1-1 mediation), the influence of a higher-level predictor on a lower-level outcome can be mediated by an intervening variable at the lower level of analysis—e.g., the effect of development methodology use (team-level) on developer performance (individual-level) being mediated by role clarity (individual-level). Much work on cross-level mediation has focused on two-level models. However, cross-level mediation can also be modeled in three-level models. For instance, the effect of a level-3 predictor on a level-1 outcome might be mediated by a level-3 mediator (as in 3-3-1 mediation) or a level-1 mediator (as in 3-1-1 mediation).[2] Additional mediators might also be incorporated at the intervening level (level-2). Further sophistication can be added when considering cross-level moderated mediation. That is, the extent to which a cross-level relationship between a level-2 factor and level-1 outcome is mediated (in either 2-2-1 or 2-1-1 mediation) can vary across level-3 units. Procedures for testing cross-level mediation in such models differ depending on the level of analysis at which the mediator resides. Specifically, when the mediator resides at lower levels of analysis (e.g., level-1 and level-2 in a three-level context) non-independence within higher-level units needs to be taken into account (Bauer et al., 2006). Such considerations are not necessary when the mediator resides at the highest level (e.g., when the mediator resides at level-3 in a three-level model). Although these various forms of cross-level mediation have potential for multilevel theory development and testing, detailed discussion of them is beyond the scope of this research.

Guided by the bridging principle where hierarchically distal factors are linked by mechanisms at an intervening level of analysis (Hackman, 2003; Rousseau, 2011), we focus on situations where the predictor, mediator, and outcome reside at different levels of analysis. From a theoretical perspective it is important to ensure that there is sufficient bond strength across each level of analysis (Klein & Kozlowski, 2000). Heeding bond strength across levels of analysis ensures that theoretically relevant mechanisms through which higher-level factors affect lower-

[2] Note that although such cross-level mediation models can be statistically tested, they run the theoretical risk of overlooking important mechanisms at the intervening level of analysis (level-2). IS researchers need to give careful consideration to the theoretical impetus for examining such models (Hackman, 2003).

level outcomes are not missed by skipping intervening levels of analysis (Klein & Kozlowski, 2000). For instance, in their examination of the cross-level effects of empowerment climate, Seibert et al. (2004) posited that empowerment climate at the business unit level would affect individual job performance through its influence on individual psychological empowerment. According to the bond strength principle, although empowerment climate resides at a higher-level of analysis, it is a theoretically proximal predictor of psychological empowerment (i.e., it is reasonable to expect that individuals embedded within an environment that promotes empowerment will experience higher levels of psychological empowerment compared to individuals embedded in an environment that does not promote such empowerment), which in turn is theoretically proximal to individual job performance. Therefore, predictors identified at level-3 must be reasonably proximal (from a theoretical standpoint) to theoretical entities at level-2. Similarly, predictors at level-2 must be reasonably proximal to theoretical entities at level-1. Consideration of bond strength ensures that researchers identify and develop sound meditational mechanisms in their theories. Figure 6 shows variations in the form that cross-level mediation can take in three-level models.

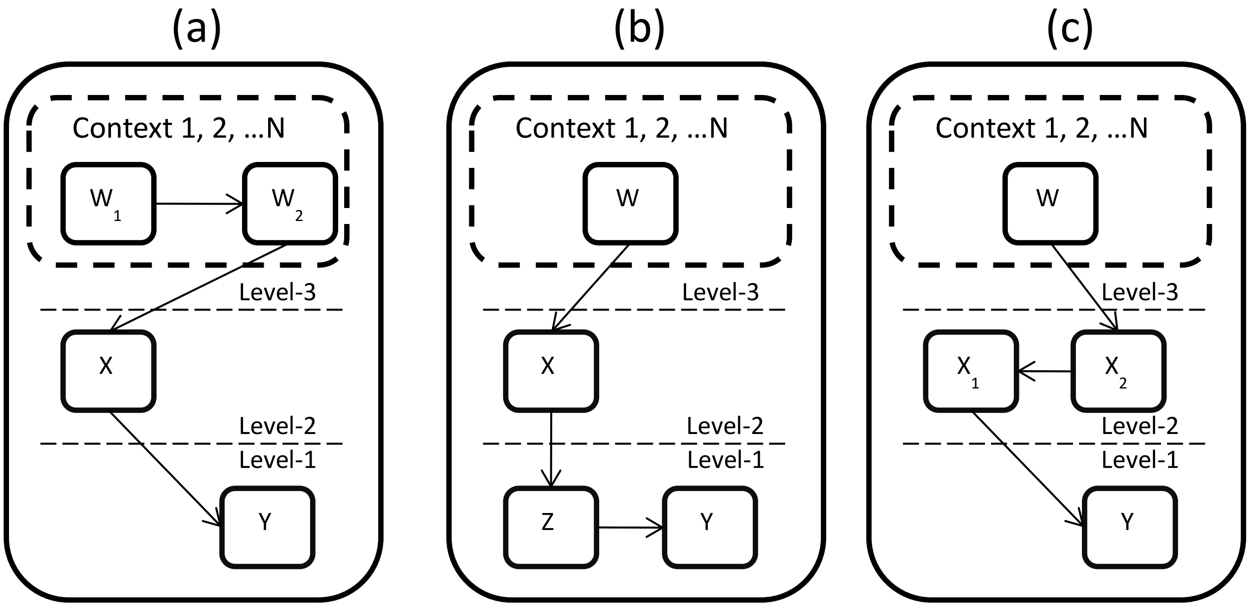


Figure 6. Different Forms of Cross-Level Mediation in Three-Level Models

Once the theoretical bond strength of constructs at different levels of analysis has been assessed, the researcher can proceed with testing for cross-level mediation. We build on Mathieu and Taylor's (2007) framework for meso-mediation testing, given its focus on linking multilevel theory and testing. Although their framework was developed in a two-level context, it can be extended to a three-level context using a similar blueprint (Mathieu & Taylor, 2007; Raudenbush & Bryk, 2002). Mathieu and Taylor's (2007) meso-mediational framework builds heavily on Baron and Kenny's (1986) test for mediation. In short, this process requires that: (1) there should be a significant relationship between the independent variable (X) and dependent variable (Y); (2) there should be a significant relationship between the independent variable (X) and the mediator variable (M); (3) there should be a significant relationship between the mediator variable (M) and the dependent variable (Y); and (4) the previously significant relationship between the dependent (Y) and independent variables (X) should become non-significant (or weaker) after controlling for the mediator variable (M). However, given the nested structure of multilevel data and the associated non-independence of observations, there are some additional complexities involved in cross-level mediation testing (Bauer et al., 2006; Krull & MacKinnon, 2001). Specifically, testing of lower-level relationships needs to account for variance across higher-level units (Bauer et al., 2006). This is where two-level models fall short, as they are unable to account for variance at level-3 when examining cross-level mediation from level-2 to level-1. Thus, a 3-3-2-1 cross-level mediation model, such as the one presented in Figure 6a, cannot simply be broken down into two two-level models (e.g., mediation from level-3 to level-2 and mediation of level-2 to level-1) because, as noted earlier, the level of the predictor and the mediator is a major consideration in the evaluation of cross-level mediation models. An integrated three-level model avoids biased estimates by accounting for non-independence of level-1 and level-2 units simultaneously.

The first step in cross-level mediation testing is to determine where the variance in the lower-level outcome resides, per Guideline 1. Clearly, there would be no need to pursue three-level mediation testing if level-3 unit membership did not account for any variance in the lower-level outcome of interest. Assuming there is between-unit variability at level-3, the next step is for the researcher to determine if there is sufficient variability in potential lower-level mediators that can be accounted for by higher-level unit membership. Building on the developer performance example, suppose project team process (level-2) is posited to mediate the relationship between client-firm characteristics (level-3) and developer performance (level-1). The researcher would need to demonstrate that some proportion of the variance in project team process is attributable to client firm membership (i.e., to what extent does project team process vary as a function of which client firm a project team is assigned to?). This can be accomplished with a simple unconditional two-level model with project team process as the lower-level outcome. The intraclass correlation coefficient (ICC) would indicate how much of the variability in project team process is accounted for by client firm membership and a χ^2 test will indicate whether this variability is statistically significant.

Mathieu and Taylor (2007) recommended that researchers begin by testing within-level relationships before testing cross-level relationships as these might also have cross-level effects. In the interest of simplicity, we do not include within-level mediation relationships. However, procedures for including such relationships are enumerated in Mathieu and Taylor (2007). The next steps are fairly similar to Baron and Kenny's (1986) guidelines for mediation testing, but with some subtle, yet important, differences. Specifically, using a three-level model, one regresses the outcome—developer performance (level-1)—on the higher-level predictor—client firm characteristics (level-3)—i.e., the X→Y relationship. Next, the mediator—project team process (level-2)—is regressed on the level-3 predictor client firm characteristics—i.e., the X→M relationship. This can be accomplished using a two-level model using project team process as the outcome. One then regresses developer performance on the mediator—project team process—i.e., the M→Y relationship. However, as noted earlier, the researcher needs to be careful to evaluate this step using a three-level model because two-level models fail to account for non-independence due to nesting of this relationship within level-3 units. Importantly, the strength of the cross-level relationship between project team process and developer performance could vary as a function of client firm membership. Simply decomposing the mediation testing into two two-level models does not account for this non-independence and can lead to biased estimates (Bauer et al., 2006). Finally, developer performance is regressed on client firm characteristics and project team process—i.e., X→M→Y—to determine if the effects of the predictor on the outcome are partially or fully mediated. A Sobel test can then be conducted to determine the extent to which the effects of the level-3 predictor on the level-1 outcome are carried, indirectly, through the level-2 mediator (Krull & MacKinnon, 2001; MacKinnon et al., 2002; Mathieu & Taylor, 2007).[3] Table 3 below provides a summary of the steps involved in cross-level mediation testing using three-level models.

Table 3. Summary of Guidelines for Cross-Level Mediation in Three-Level Models

Guideline	Objective	Steps
Guideline 1 (Variance components)	Determine where variability in lower-level outcome resides	• Step 1: Use a three-level unconditional model to decompose variance into three-level components
Guideline 2.3 (Cross-level mediation)	If mediator resides at highest level of analysis (i.e., level-3)	• Step 1: Run a three-level conditional model using level-3 predictor (excluding level-3 mediator) • Step 2: Run a simple ordinary least squares (OLS) regression model to test the relationship between the level-3 predictor and the level-3 mediator • Step 3: Run a three-level model to test the relationship between the level-3 mediator and level-1 outcome while accounting for level-2 unit membership • Step 4: Run a three-level model including the level-3 antecedent and the level-3 mediator predicting the level-1 outcome • Step 5: Use a Sobel test to assess the significance of indirect cross-level effect
	If mediator resides at lower level, determine the extent to which the lower-level mediator varies as a function of higher-level unit membership	• Step 1: Follow Guideline 1 to determine where variance in outcome variable resides • Step 2: Run a two-level unconditional model using level-2 mediator as the lower-level outcome
	Determine whether the higher-level antecedent affects the lower-level outcome	• Step 1: Run a three-level conditional model using level-3 predictor (excluding level-2 mediator) • Step 2: Run a two-level conditional model using the level-3 variable as a predictor for the level-2 mediator

[3] Recent literature argues that, while useful, the guidelines outlined by Baron and Kenny (1986) to test for mediation are quite liberal and that more rigorous tests can provide information that complements the insights yielded by the Baron and Kenny approach (MacKinnon et al., 2002; Mathieu & Taylor, 2007). The Sobel test has been advocated as a more rigorous test for determining the extent to which the effects of a predictor on the outcome are carried through the mediator (MacKinnon et al., 2002). The Sobel test is robust to deviations from normality in the distribution of residuals (MacKinnon et al., 2002).

		• Step 3: Run a three-level model to test the relationship between the level-2 mediator and level-1 outcome while accounting for level-3 unit membership • Step 4: Run a three-level model including the level-3 antecedent and the level-2 mediator predicting the level-1 outcome • Step 5: Use a Sobel test to assess the significance of indirect cross-level effect

4.2.4 Guideline 2.4: Constructing Three-Level Moderated Mediation and Mediated Moderation Models

Besides the three-level models discussed above (i.e., main effects, cross-level moderation, cross-level mediation), more sophisticated three-level models can include moderated mediation and mediated moderation. We graphically illustrate some potential variations of these models in Figure 7. Specifically, Figures 7a and 7b are two potential cross-level moderated mediation three-level models, with a moderator from level-3 and level-2, respectively. Similarly, Figures 7c and 7d are two potential cross-level mediated moderation three-level models, with a moderator from level-3 and level-2, respectively. If these sophisticated models are appealing for theoretical reasons, researchers can leverage and use them. Using these models calls for combining the three-level guidelines previously presented and established guidelines about testing direct and indirect effects (e.g., Bauer et al., 2006; Preacher, 2011). Models with such sophistication are still fairly uncommon. We found no paper using these models in our literature review of IS papers. We extended our search to major management journals[4] in the last five years and found only one paper (i.e., Ou et al., 2017) that incorporates mediated moderation and moderation mediation for three-level data (i.e., individual, leader, and company).

[4] Journals we reviewed include *Academy of Management Journal, Academy of Management Review, Administrative Science Quarterly, Journal of Applied Psychology, Journal of Management, Organization Science, Personnel Psychology, Strategic Management Journal,* and *Organizational Behavior and Human Decision Processes.* The goal here is not conducting an exhaustive literature review but to identify articles to provide examples for researchers as references. We therefore focused on teams research in the last 5 years.

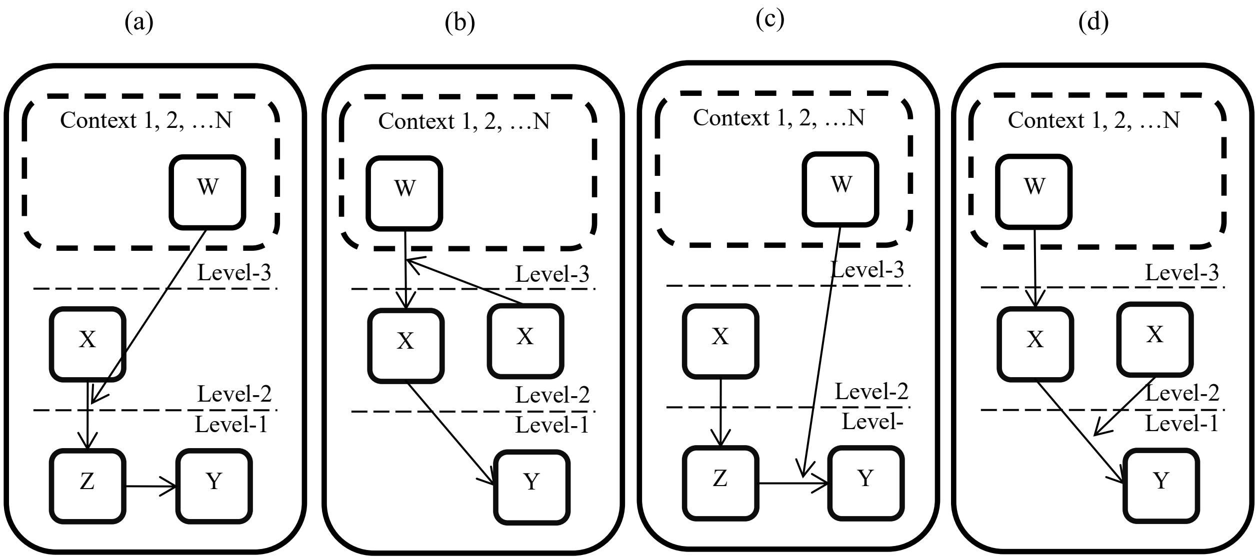


Figure 7. Examples of Cross-Level Moderated Mediation and Mediated Moderation in Three-Level Models

4.2.5 Guideline 2.5: Context as Cross-Classification

We have thus far discussed how studies can represent relationships that cross three levels and that are hierarchically nested. However, researchers may come across situations in which entities at different levels are not hierarchically nested but entities at the lowest level can be meaningfully grouped within combinations of entities at the two higher levels. Such a classification structure may be caused by crossed organizational membership, as in multiple membership models, or by groups that change in membership over time, as in dynamic group models.

Cross-classified models can be used when the three levels are not hierarchically nested, but when observations at level-1 can be grouped using combinations of level-2 and level-3 groups.[5] Researchers can employ cross-classification models to more precisely conceptualize and effectively represent the nature of the context in which level-1 outcomes and explanatory

[5] For technical aspects of the model specification of the modeling techniques, see Cafri et al. (2015). Other resources include Shi et al. (2010) for cross-classified models, and Fielding and Goldstein (2006) for cross-classification and multiple membership models.

mechanisms function. Overlooking this cross-classification would lead to an underspecified model, biased estimates, and incorrect partitioning of variance of the outcomes (Shi et al., 2010). Conversely, incorporating cross-classification to capture the contextual embedding of level-1 observations can be useful to improve the quality of the estimates of explanatory effects, identify components of variance in outcomes, and study differential effects of cross-classifications.

We discuss here how different contexts can be captured using specific types of cross-classified models. Specifically, we consider plural association with groups, which we call "context as plural association", and association with dynamic groups, which we call "context as dynamic entities". First, *multiple membership models* can be used to model "*context as plural association"*, in which observations at level-1 can simultaneously be members of more than one level-2 group (or level-2 observations can be simultaneously members of more than one level-3 groups). A key issue for multiple membership models is to decide how much weight to place on membership in each group. Second, *dynamic group models* can be used to model "context as dynamic entities", in which higher-level groups are dynamic entities, i.e., when groups change over time (Bauer et al., 2013). In such a context, assuming groups are static and overlooking the changing nature of groups over time may result in an incorrect model specification, because, when groups change, their effects may change as well. Moreover, inappropriately modeling time-varying group effects may overestimate the standard error of cross-level interactions, resulting in decreased power to detect effects (Cafri et al., 2015). As such, the changing effects of groups represent a dynamic context for lower-level entities in which their outcomes and explanatory mechanisms function.

4.2.6 Guideline 2.6: Conducting Robust Three-Level Analysis

Like with other methods, analyses of three-level models need to address potential endogeneity concerns. Where the researcher seeks to estimate the "within" effect (i.e., the effect of a level-1 regressor on a level-1 outcome that is typically of interest and relevant for causal interpretation), failure to correctly model the unobserved variation due to the hierarchical structure of the data can introduce endogeneity into the model (Antonakis et al., 2021). As a basic rule, level-1 outcomes should not be driven by factors (observable or unobservable) other than the regressors. As a result, unaddressed endogeneity may lead to underestimated or overestimated models. Major sources of endogeneity include omitted variables (e.g., temperature is not included when predicting ice cream price, but the vendor increases ice cream price in spring), simultaneity (e.g., coupons cause more purchase and customers who purchase more receive more coupons), and selection bias (e.g., sampling only productive researchers when studying research productivity) (Heckman, 1979).

Researchers need to consider new potential threats to endogeneity when using three-level models. A key assumption in multilevel models is that the unobserved random effects are uncorrelated with level-1 regressors (Raudenbush & Bryk, 2002, p. 255). Accordingly, endogeneity in 2-level models can arise when the assumption that level-2 error terms are uncorrelated with level-1 regressors is violated, resulting in the failure of the random effects assumption (Antonakis et al., 2021). By extension, endogeneity in 3-level models can arise when the assumption that higher-level (level-3, level-2) error terms are uncorrelated with lower-level (level-2, level-1) regressors is violated. Omitted variables at level-2 and level-3 may contribute to systematic differences in level-1 regressors and outcomes, and overlooking these differences can lead to endogeneity problems. Rabe-Hesketh and Skrondal (2021) provided a classification of endogeneity issues and the corresponding solutions for two-level models (p. 270), which can similarly be applied to three-level models. Additionally, Antonakis et al. (2021) offered a set of practical recommendations for organizational researchers to model multilevel data appropriately to avoid endogeneity.

Addressing endogeneity concerns in three-level models requires similar strategies employed in OLS regression analysis, with some additional complexity. For example, endogeneity due to omitted variables can be potentially addressed by including instrumental variables (i.e., in the ice cream example, temperature can serve as the instrumental variable). With multilevel analysis, instrumental variables from the same or higher levels can be used for lower-level predictors. For example, Wang et al. (2019) examined the role of opportunity contexts in a longitudinal study of unauthorized access attempts on informational systems. In their three-level analysis that predicted the dependent variable of unauthorized access attempts, a department-level (level-3) variable of off-hour-needed, measured using a human resources manager's evaluation on each department's "need to work outside of the regular business hours" (p. 615), is used as the instrumental variable for the observation level (level-1) predictor off-hour access, and an employee-level (level-2) variable of internet-enabled, measured using "the percentage of internet-accessible applications in an employee's system profile" (p. 616), is used as the instrumental variable for the observation-level (level-1) predictor of off-site access. Endogeneity due to simultaneity may be harder to address. A commonly used technique is including lagged variables. Instead, three-level models can be used to incorporate the time dimension for a two-level model, i.e., recording observations nested in groups longitudinally. Researchers can also use a blend of a mixed model and a Heckman model and then compare the model fit. For example, Xie and Lee (2015) employed both lagged variables and Heckman models in their study on the effects of exposure to social media on purchase behavior. Overall, addressing endogeneity is closely related to the research context and theoretical grounding.

4.2.7 Guideline 2.7: Centering Techniques in Three-Level Analysis

An important technique for multilevel modeling is the centering methods. Centering is a technique that rescales predictors by subtracting their means. Centering in multilevel models involves two types of means: group mean and grand mean, with the former referring to the mean of each group and the latter the mean of the full sample. Whether to use grand mean centering or group mean centering does not impact significance tests and is mostly for the purpose of improving interpretation. Lower-level variables have the options of both group mean centering and grand mean centering, but the highest level (level-3) only has the option of grand mean centering. Centered means can be reintroduced at higher levels if the researcher wishes to investigate separate within-group and between-group effects of the predictor (Raudenbush & Bryk, 2002). We also recommend researchers consult existing discussions on this topic (e.g., Algina & Swaminathan, 2011; Enders & Tofighi, 2007; Raudenbush & Bryk, 2002; Snijders & Bosker, 2011). Overall, the decision regarding non-centering, grand mean centering, or group mean centering depends on the research question of interest and the research context.

The guidelines outlined above provide researchers with a useful set of steps to consider when connecting multilevel theory with empirical testing. They are particularly useful for two multilevel theoretical goals—(1) incorporating discrete context at multiple levels by modeling variables that form a part of the higher-level context and (2) bridging constructs across hierarchically distal levels of analysis. It is prudent to emphasize that multilevel modeling is not intended to serve as a "hammer" for examining all phenomena in IS research. The determination of whether multilevel investigation is warranted should be based on the theoretical objectives of the researcher, the hierarchically nested structure of the study setting and data collection. The researcher must consider such issues as whether it makes sense, theoretically, that higher-level factors might exert a major influence on the lower-level phenomenon of interest, and just how strong higher-level influences are likely to be in the study context. Following Guideline 1 should help IS researchers in making this determination. Next, we illustrate the utility of these guidelines in the context of an empirical study of offshore software development projects.

5 Illustration

One domain in which hierarchical nesting of units is common is software development. Given the complexity and knowledge-intensive nature of software, successful project completion requires the input of multiple software developers working interdependently to construct a product that meets client needs (Walz et al., 1993; Warkentin et al., 2009). In these settings, software developers are nested within software project teams (Faraj & Sproull, 2000; Maruping et al., 2009). Further, these software project teams, when housed within software vendors, work for client firms, and interact with client liaisons who are responsible for coordinating key project milestones (Westner & Strahringer, 2010). Client liaisons facilitate communications between client firms and software project teams, each responsible for managing the projects for a client firm. Vendors typically assign a client liaison to manage the interactions with a particular client. As an example, a client may have had three customized software solutions developed by a vendor, with their development assigned to three different software project teams. The client liaison would serve as a key contact point among the vendor, the project teams, and the client organization to ensure the project teams understand and meet the client's needs and expectations. Thus, software project teams are nested within client liaisons. In this multilevel nested structure, at level 1 are software developers, at level 2 are project teams, and at level 3 are client liaisons.

In this particular illustration, we examined software developers who were embedded in offshore software project teams. The offshore software project teams completed projects for client firms. The main role of a client liaison was to see to the completion of the project to the satisfaction of the client. In this context, we were interested in understanding the link between client liaison characteristics and their impact on project outcomes. Although the extant literature has examined project leadership and its effect on software project team performance (e.g., Faraj & Sambamurthy, 2006; Maruping et al., 2009), the link between client liaison leadership characteristics and individual developer outcomes is much less well-studied. We found this to be particularly important given the central role that developers play in creating software functionality (Fang & Neufeld, 2009; Walz et al., 1993) and the ability of client liaisons to create an environment that influences developers (Gopal & Gosain, 2010; Herbsleb & Mockus, 2003; Levina & Vaast, 2008). The few studies that examine cross-level relationships among client firm, project team, and developer in offshore ISD projects are limited to two-level models: between client firm level leadership and project team (e.g., Rai et al., 2009) or between project team and developer (Windeler et al., 2017). A three-level approach has the potential to synthesize these relationships and shed light on the mechanisms bridging client liaison level characteristics and individual developer outcomes through the project team-level (meso-level) structures that they shape. Specifically, we examine two different characteristics of client liaisons using three-level models. The first model focuses on their experience and the second focuses on their transformative leadership. Figures 8a and 8b show the two alternative models in which we are interested.

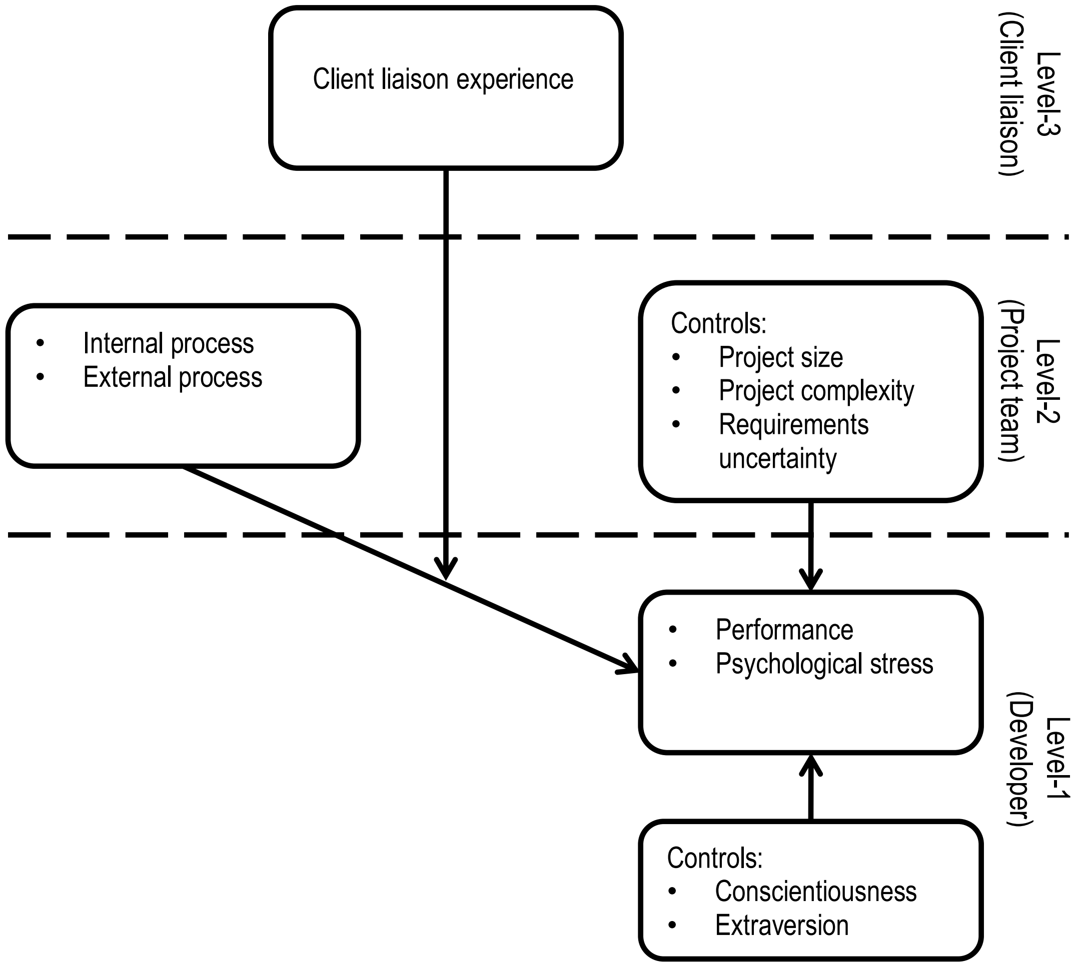


Figure 8a. Cross-Level Moderation Model of Client Liaison Experience, Team Process, and Developer Outcomes

5.1 Cross-Level Moderation Model

As shown in Figure 8a, we wanted to understand the cross-level moderating role of client liaison leadership so as to shed light on how elements of discrete context affect the nature of lower-level cross-level relationships. In this case, we were interested in examining the extent to which the cross-level effects of internal and external team processes on individual developer performance and psychological stress vary as a function of the client liaison leader to whom the project team was assigned. We expected that project teams implementing internal and external processes would benefit more from these processes when they had an experienced client liaison leader compared to project teams assigned to less experienced client liaison leaders. The rationale is that experienced client liaison leaders are able to provide better guidance to developers on how to translate collective processes into individual effectiveness (Hackman & Wageman, 2005). We expect such client liaisons to be better aware of strategies involved in interfacing with clients from another culture and to provide insights that give greater clarity on client expectations. Thus, incorporating this element of the context within which offshore project teams were embedded would shed light on the nature of the cross-level relationship between team processes and individual developer outcomes.

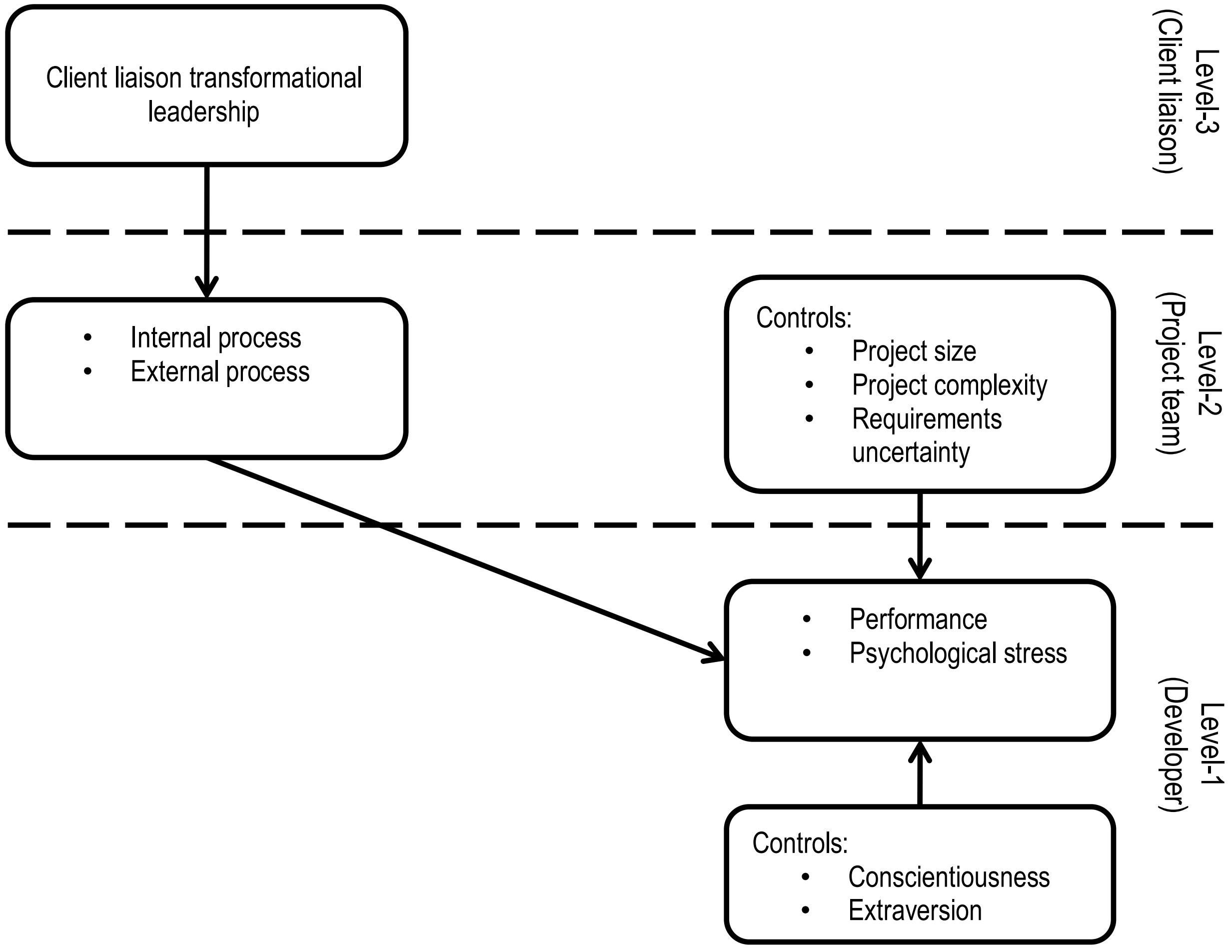


Figure 8b. Cross-Level Mediation Model of Client Liaison Transformational Leadership, Team Process, and Developer Outcomes

5.2 Cross-Level Mediation Model

As shown in Figure 8b, we also sought to bridge the client liaison level with individual developer outcomes. Rai et al. (2009) found that project leader cultural values influenced offshore project outcomes. However, they neither identified the mechanisms through which this happens—i.e., the project team-level structures that such leaders created for their teams to achieve success—nor examined the implications for individual developers. We felt it was important to examine these mechanisms because activities that benefit project teams as a unit are not necessarily beneficial to individual team members (Marrone et al., 2007). In this particular example, we examine the possibility that client liaisons who exhibit *transformational leadership* values are likely to institute both integrative and autonomous team processes (Jiang & Chen, 2018). Internal team processes—which involve activities directed at the establishment of standardized procedures for determining who will take ownership over specific tasks, how resources will be deployed, and how progress toward project goals will be monitored (Marks et al., 2001)—and external team processes—which represent the means by which teams go beyond their boundaries to interact with external constituents (Ancona, 1990)—constitute such activities (Ancona & Caldwell, 1992; Marks et al., 2001). Jiang and Chen (2018) found that transformational leaders placed a stronger emphasis on instituting team cooperative norms to facilitate knowledge sharing and autonomous culture to promote innovation, suggesting that client liaison leaders embracing such values are highly likely to enforce the use of internal team processes. Additionally, external team processes—particularly those aimed at interacting with the client—enable client liaison leaders to get more information about the client (Gopal & Gosain, 2010; Rai et al., 2009). Hence, we expected that client liaisons who exhibit transformational leadership values would promote the use of such activities in the project teams for which they were responsible and that these activities would have implications for the performance and well-being for project team members.

Next, we briefly describe the study setting and data collected for this illustration. We then illustrate the application of the guidelines outlined earlier.

5.3 Sample and Participants

We conducted a field study of offshore IT projects managed by a leading software vendor in India and several of its U.S.-based clients. Our sampling frame was 234 agile IT projects, completed in 2017, across 26 client firms, with each client firm having at least 5 projects. Each project was completed by a team of 8-12 members. All of the projects were custom developed by the vendor to the specific requirements of each U.S. client on respective projects, and the projects developed for each client were overseen by a particular client liaison. Examples of projects included a complete human resource management system and a customer relationship management system. There were a total of 234 different project teams, with 2,302 team members across all the projects. In the interest of brevity, details regarding the measurement are included in Appendix B. Table 4 provides a summary of the variables involved in our analysis and Table 5 shows the descriptive statistics and correlations between these variables.

Table 4. List of Variables

Variable	Source
Individual level	
Developer performance	Project archives
Psychological stress	Developer survey response
Team level	
Internal team process	Developer survey response
External team process	Developer survey response
Client liaison leadership level	
Experience	Organization archives
Transformational leadership	Developer survey response
Individual-level controls	
Conscientiousness	Developer survey response
Extraversion	Developer survey response
Team-level controls	
Project size	Project archives
Project complexity	Project archives
Requirements uncertainty	Project archives

Table 5. Correlations, Means, and Standard Deviations

	Mean	SD	1	2	3	4	5	6	7	8	9	10
1. Performance	4.98	1.04										
2. Psychological stress	5.07	1.17	-.21**									
3. Project complexity	8,987	2,134	-.21**	.28***								
4. Requirements uncertainty	17.66	7.13	-.25***	.24***	.31***							
5. Project size	410,040	56,585	-.25***	.29***	.10	.17**						
6. Client liaison experience	7.35	2.15	.24***	-.19**	.10	.19**	.23***					
7. Client liaison transformational leadership	5.06	1.81	.19**	.21**	.09	.11*	.14*	.10				
8. Internal process	3.95	1.31	.25***	.29***	.15*	.19**	.17**	.19**	.24***			
9. External process	4.07	1.46	.34***	.35***	.14*	.17**	.20**	.21***	.24***	.24***		
10. Conscientiousness	5.55	0.75	.16**	.14*	.11*	.08	.10	.17**	.11*	.14*	.11*	
11. Extraversion	3.17	1.50	.10	.19**	.09	.10	.14*	.19**	.13*	.19**	.08	.17**

Notes:
1. Level-1, n=2,302; Level-2, n=234; Level-3, n=26.
2. * p < .05; ** p < .01; *** p < .001.

5.4 Application of the Guidelines

Given the hierarchically nested structure of the data and the multilevel focus of the models being tested, we used HLM8 (Raudenbush et al., 2004)—a RCM software package—to conduct our analyses. HLM is particularly well-suited for estimating three-level models.

5.4.1 Application of Guideline 1

Guideline 1 suggests that researchers use theory and analysis of the variance components to determine the appropriateness or necessity for three-level models. In this particular context, the theoretical model was aimed at incorporating discrete context, thus, providing a theoretical rationale for pursuing a three-level model. The study context also provided an empirical rationale for considering a three-level model. Specifically, because offshore software project teams were nested within client liaisons, it was highly likely that they were exposed to common within-client liaison influences. For example, the same client liaison may promote similar behaviors across different software project teams such as scheduling, communication, and expectation management, that in turn may affect both project and individual outcomes of interest. Given this rationale, the next step was to determine whether there was sufficient variability in the outcomes of interest—in this case, individual developer performance and psychological stress—to warrant a three-level model. To accomplish this, we estimated a three-level unconditional model to obtain the variance components for developer performance and psychological stress across the three levels of analysis. The system of equations is outlined in Appendix A. As Table 6 shows, the results of a χ^2 test indicate that there was sufficient variability at all three levels of analysis. Specifically, 54.5% of the variance in individual developer performance was attributable to differences between individuals; 33% ($p < .001$) of the variance was due to between-project team differences and 24.6% ($p < .001$) of the variance was due to between-client liaison differences. In the case of individual psychological stress, 53.5% of the variance was between individuals, 28.9% ($p < .001$) was between project teams, and 26.4% ($p < .001$) was between client liaisons. These results suggest that it was appropriate to proceed to identify predictors at these different levels of analysis that explain the variance in these dependent variables.

Table 6. Three-Level Unconditional Model Predicting Developer Performance and Psychological Stress

Variance component	Variance estimate	χ^2	p-value
Developer performance			
Level-1 variance (e_{ijk})	1.3365		
Level-2 variance (r_{0jk})	.8788	2986.50	0.00
Level-3 variance (u_{00k})	.6321	1012.58	0.00
Psychological stress			
Level-1 variance (e_{ijk})	1.3169		
Level-2 variance (r_{0jk})	.7969	3601.65	0.00
Level-3 variance (u_{00k})	.6540	1380.65	0.00

Notes:
1. Level-1, n=2,302; Level-2, n=234; Level-3, n=16.
2. *** $p < .001$.

5.4.2 Application of Guideline 2.2

Because the results of following Guideline 1 suggested that a three-level model was appropriate, we proceeded to follow Guideline 2.2 for cross-level moderation. The guideline suggests that the cross-level effect of team processes (level-2) on individual outcomes (level-1) should be examined while controlling for level-3 unit membership. Table 7 (model 2) shows the results of the cross-level main effects. Specifically, internal and external team process had a significant cross-level effect on developer performance (internal team process: $\gamma = .27$, $p < .001$; external team process: $\gamma = .24$, $p < .001$) and psychological stress (internal team process: $\gamma = .20$, $p < .001$; external team process: $\gamma = .15$, $p < .01$). In this model, the coefficients for internal and external team processes were treated as fixed effects. In model 3, the coefficients for internal and external team processes were allowed to vary randomly. As the deviance statistics suggest, this yielded a better model fit (developer performance: $\text{deviance}_{\text{fixed}}$ = 11770.65 versus $\text{deviance}_{\text{random}}$ = 11522.68; psychological stress: $\text{deviance}_{\text{fixed}}$ = 11987.68 versus $\text{deviance}_{\text{random}}$ = 111526.45). A likelihood ratio test indicated that this constituted a significant improvement over the model where these effects were fixed.

The results of model 3 show that the slopes (coefficients) for internal team process (U_{04k} = .14, $p < .01$) and external team process (U_{05k} = .13, $p < .01$) varied significantly across client liaisons in predicting developer performance. Similarly, the slope variance was significant for internal team process (U_{04k} = .15, $p < .01$) in predicting psychological stress. These results suggest that the cross-level effects of team processes on developer performance and psychological stress varied across client liaisons and that we should examine the effects of potential moderators at the client liaison level (level-3). Hence, we proceeded to include the client liaison-level moderator. Model 4 shows the results of the cross-level moderating effect of client liaison experience on the cross-level relationship between team processes and individual outcomes. As the results indicate, client liaison experience had a significant cross-level moderating effect on the relationship between team processes and developer performance (client liaison experience X internal team process: $\gamma = .15$, $p < .05$; client liaison experience X external team process: $\gamma = .16$, $p < .05$) and psychological stress (client liaison experience X internal team process: $\gamma = -.13$, $p < .05$). The decrease in the deviance statistics[6] suggests a significant improvement to the model fit (developer performance: $\text{deviance}_{\text{main effects}}$ = 11522.68 versus $\text{deviance}_{\text{interaction}}$ = 11136.87; psychological stress: $\text{deviance}_{\text{main effects}}$ = 111526.45 versus

[6] The deviance statistic is the best indicator of model fit in RCM (Raudenbush & Bryk, 2002). It is also possible to compute a Pseudo-R^2 that is computed as a ratio of total variance from an unconditional model and unexplained variance from a conditional model (Snijders & Bosker, 1999). However, caution is urged regarding the Pseudo-R^2 as it can be unstable and has the potential to under- or over-estimate true effect sizes (Snijders & Bosker, 1999). An alternative approach is to use the cross-level operator analysis (James & Williams, 2000) to model the total variance. The variance explained statistic produced by this approach is similar to a traditional R^2 from regression analysis. This approach yields a better estimate of the variance explained by the predictive model. We report this total R^2 in our results.

deviance$_{interaction}$ = 10999.87). These results show that the cross-level relationship between team processes and developer performance and psychological stress vary as a function of client liaison experience. The interaction plots shown in Figure 9(a, b) show that internal and external team processes had a stronger positive cross-level effect on individual developer performance in projects handled by more experienced client liaisons compared to projects handled by less experienced client liaisons. Figure 9c shows that internal team processes had a stronger negative cross-level effect on individual psychological stress in projects, with more experienced client liaisons compared to projects handled by less experienced client liaisons. The system of equations used to test for cross-level moderation is included in the Appendix A.

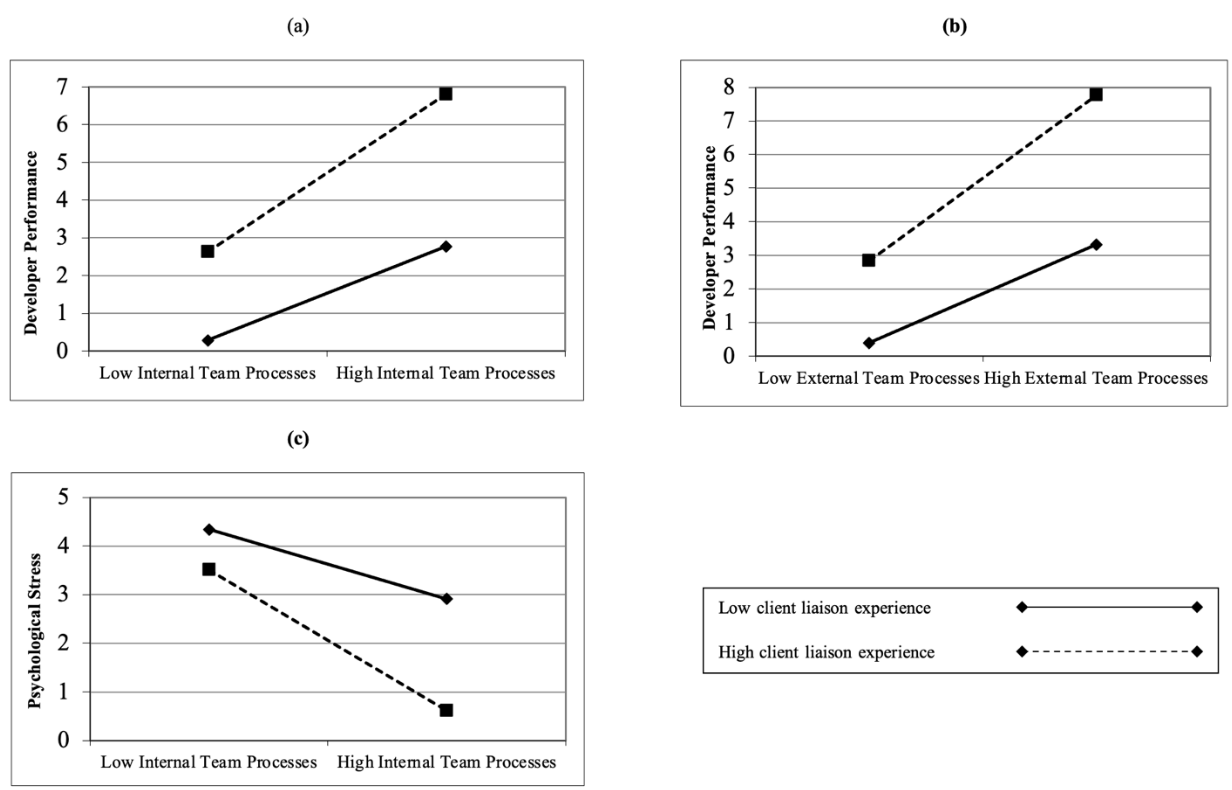


Figure 9. Plots of Cross-Level Interactions Between Client Liaison Leader Experience and Team Processes

Table 7. Three-Level Cross-Level Moderation Model Predicting Developer Performance and Psychological Stress

	Performance				Psychological stress			
Variable	1a	2a	3a	4a	1b	2b	3b	4b
Level-1 (Main effects):								
Intercept (π_{0jk})	.19 (.054)	.17 (.040)	.14 (.042)	.13 (.055)	.15 (.040)	.13 (.044)	.12 (.059)	.10 (.074)
Conscientiousness (π_{1jk})	.14* (.013)	.08 (.022)	.06 (.031)	.05 (.040)	.08 (.010)	.07 (.024)	.05 (.027)	.03 (.030)
Extraversion (π_{2jk})	.06 (.020)	.06 (.024)	.05 (.032)	.04 (.030)	.06 (.015)	.05 (.010)	.04 (.019)	.03 (.022)
Level-2 (Main effects):								
Project size (β_{01k})	-.14* (.010)	-.14 (.071)	-.10 (.032)	-.08 (.080)	.14** (.012)	.12* (.015)	.12* (.016)	.12* (.017)
Project complexity (β_{02k})	-.12* (.014)	-.10 (.038)	-.09 (.041)	-.05 (.050)	.18*** (.010)	.16** (.012)	.15** (.019)	.14* (.020)
Requirements uncertainty (β_{03k})	-.14* (.011)	-.07 (.041)	-.06 (.050)	-.05 (.055)	.17** (.011)	.14** (.013)	.14** (.014)	.13* (.014)
Internal team process (β_{04k})		.27*** (.011)	.21*** (.014)	.17** (.015)		.20*** (.010)	.18** (.012)	.13** (.013)
External team process (β_{05k})		.24*** (.012)	.22*** (.013)	.17** (.017)		.15** (.011)	.13** (.012)	.12* (.013)
Level-3 (Main effects):								
Client liaison experience (γ_{001})				.15** (.012)				.15** (.012)
(Cross-level interactions):								
Client liaison experience x internal team process (γ_{041})				.15** (.013)				-.13* (.010)
Client liaison experience x external team process (γ_{051})				.16** (.014)				.09 (.018)
Random Effects								
Level-1 variance (e_{ijk})	.30***	.27***	.23***	.19***	.31***	.29***	.26***	.23***
Level-2 variance (r_{0jk})	.21***	.20***	.17***	.13***	.23***	.21***	.19***	.14***
Level-2 slope variance (U_{04k}) (internal team process)			.14** (.010)	.12** (.011)			.15** (.019)	.06 (.033)
Level-2 slope variance (U_{05k}) (external team process)			.13** (.014)	.10* (.015)			.10 (.061)	.14* (.013)
Level-3 variance (U_{00k})	.06***	.06***	.05***	.03*	.06**	.06***	.04***	.03***
Deviance	11946.51	11770.65	11522.68	11136.87	12622.84	11987.68	111526.45	10999.87
χ^2	10927.64	9984.65	7966.87	6012.65	9434.60	7768.75	5940.31	1769.85
R^2	.10	.17	.22	.26	.12	.16	.20	.26

Notes:
1. Level-1, n=2,302; Level-2, n=234; Level-3, n=26.
2. * p < .05; ** p < .01; *** p < .001.

As the results in Table 7 indicate, of all the models we compared, the three-level models (i.e., 4a, 4b) yielded the best model fit based on the deviance statistics. Thus, a focus on a two-level model of team process and developer outcomes would have led to erroneous assumptions about the nature of these cross-level relationships. Without taking client liaison leadership context into account, one might conclude that the cross-level effects of team processes on developer outcomes were uniform even when accounting for project-level differences. By including context as the third level within which project teams were embedded, we were able to gain additional insights about how the nature of these cross-level relationships differs across client liaisons. Further, the three-level model enabled us to identify client liaison experience as a discrete element of the project team context that explains why these cross-level relationships vary. This highlights the potential of three-level models to incorporate context into multilevel theory development and testing. Use of a two-level model would not have yielded such insights, as it does not examine variability in the level-2 coefficients that affect level-1 outcomes. With a two-level model, failure to account for this non-independence in level-3 units would have yielded biased estimates for the effects of team processes on developer outcomes.

5.4.3 Application of Guideline 2.3

Guideline 1 showed that variability in developer performance and psychological stress was attributable to between-project team differences as well as between-client liaison differences. Given that the mediators of interest—internal and external team processes—reside at the lower-level (level-2) and that the variability in the outcomes of interest is attributable to level-3 membership, the next step was to determine whether the mediators themselves varied as a function of client liaison membership, as suggested by Guideline 2.3. To determine this, we estimated two-level unconditional models with internal and external team processes as level-1 outcomes (see Table 8 below). The ICCs obtained from these models indicated that 43.1% (χ^2 = 401.35, $p < .001$) and 43.8% (χ^2 = 308.46, $p < .001$) of the variance in internal and external team processes respectively, were attributable to client liaison membership. These initial steps suggest that sufficient variability exists at the three levels of analysis and that it was reasonable to proceed with cross-level mediation testing.

Table 8. Two-Level Unconditional Model Predicting Internal and External Team Process

Variance component	Variance estimate	χ^2	p-value
Internal team process			
Level-1 variance (r_{ij})	1.17		
Level-2 variance (U_{0j})	1.01	401.35	< .001
External team process			
Level-1 variance (r_{ij})	1.04		
Level-2 variance (U_{0j})	.79	308.46	< .001

The first step in cross-level mediation testing is to use a three-level model to examine the relationship between the level-3 predictor and the level-1 outcome. The results of model 2 in Table 9 show that client liaison transformational leadership had a significant cross-level effect on developer performance ($\pi = -.15$, $p < .05$) and psychological stress ($\pi = .16$, $p < .01$). Second, using a two-level model, we regressed internal and external team process on client liaison transformational leadership. We found that client liaison leader transformational leadership had a positive cross-level effect on internal ($\gamma = .19$, $p < .001$) and external ($\gamma = .20$, $p < .001$) team processes (see Table 10). In the third step, we used a three-level model to regress developer performance and psychological stress on internal and external team processes while controlling for level-3 client liaison membership. Model 3 in Table 9 shows that internal and external team processes had a significant cross-level effect on developer performance (internal team process: $\gamma = .23$, $p < .001$; external team process: $\gamma = .22$, $p < .001$) and psychological stress (internal team process: $\gamma = .17$, $p < .001$; external team process: $\gamma = .18$, $p < .01$). In the fourth and final step, developer performance and psychological stress were regressed on client liaison transformational leadership and team processes. As the results in model 4 of Table 9 show, the cross-level effect of client liaison transformational leadership on developer performance ($\pi = .09$, p = ns) and psychological stress ($\pi = .10$, p = ns) became non-significant in the presence of team processes.

This pattern of results suggests that team processes fully mediated the cross-level effect of client liaison transformational leadership and developer performance and psychological stress. Results of a Sobel test showed that the effects of client liaison transformational leadership on developer outcomes were carried through internal (performance: $z = 11.93$, $p < .001$; psychological stress: $z = 14.16$, $p < .001$) and external team processes (performance: $z = 10.97$, $p < .001$; psychological stress: $z = 13.45$, $p < .001$). The system of equations for testing this cross-level mediation is included in Appendix A.

Table 9. Three-Level Cross-Level Mediation Model Predicting Developer Performance and Psychological Stress

	Performance				Psychological stress			
Variable	1a	2a	3a	4a	1b	2b	3b	4b
Level-1:								
Intercept (π_{0jk})	.15 (.068)	.13 (.077)	.12 (.081)	.11 (.088)	.13 (.044)	.10 (.046)	.08 (.047)	.05 (.048)
Conscientiousness (π_{1jk})	.13* (.012)	.08 (.035)	.05 (.041)	.03 (.044)	.09 (.016)	.07 (.017)	.05 (.018)	.04 (.019)
Extraversion (π_{2jk})	.09 (.028)	.06 (.033)	.03 (.035)	.03 (.037)	.07 (.022)	.05 (.023)	.04 (.024)	.04 (.025)
Level-2:								
Project size (β_{01k})	-.13* (.011)	-.13* (.015)	-.10 (.028)	-.06 (.035)	.17*** (.009)	.15** (.009)	.15** (.009)	.13* (.014)
Project complexity (β_{02k})	-.12* (.014)	-.12 (.015)	-.07 (.028)	-.05 (.031)	.18*** (.008)	.16** (.010)	.16*** (.011)	.13* (.013)
Requirements uncertainty (β_{03k})	-.14* (.012)	-.13*** (.017)	-.10 (.043)	-.08 (.046)	.15** (.009)	.13* (.014)	.13* (.015)	.12* (.016)
Internal team process (β_{04k})			.23*** (.015)	.21*** (.017)			.17*** (.008)	.16*** (.009)
External team process (β_{05k})			.22*** (.016)	.20*** (.018)			.18*** (.009)	.14** (.012)
Level-3:								
Client liaison transformational leadership (γ_{001})		-.15* (.018)		.09 (.043)		.16** (.015)		.10 (.041)
Random effects:								
Level-1 variance (e_{ijk})	.30***	.25***	.23***	.20***	.32***	.28***	.27***	.24***
Level-2 variance (r_{0jk})	.22***	.18***	.15***	.13***	.21***	.19***	.17***	.15***
Level-3 variance (U_{00k})	.07***	.07***	.06***	.06***	.07***	.07***	.06***	.05***
Deviance	11786.54	11542.80	11427.68	10968.75	12732.61	12012.60	11732.62	11114.66
χ^2	10941.38	10138.26	10051.63	8854.32	9162.75	8375.21	8228.65	7502.13
R^2	.07	.14	.19	.22	.11	.15	.19	.22

Notes:
1. Level-1, n=2,302; Level-2, n=234; Level-3, n=16.
2. * $p < .05$; ** $p < .01$; *** $p < .001$.

As the results in Table 9 indicate, of all the models we compared, the three-level models (i.e., 4a, 4b) yielded the best model fit based on the deviance statistics. Our use of a three-level model to bridge client liaison transformational leadership and individual developer outcomes shed important light that could not be achieved through the use of a two-level model. Specifically, by using a three-level model we were able to identify and test the role of mediating mechanisms at an intervening level of analysis—the project team level. Use of a two-level model to examine the cross-level relationship between client liaison transformational leadership and developer outcomes would have overlooked the important fact that individual developers are embedded within a project team and that client liaisons affect this collective structure. Further, a two-level model would not have enabled us to model the effect of the project team-level mediator. Using two two-level models to model the mediation would have yielded erroneous estimates of the level-2 coefficients as nesting within level-3 client liaisons would not have been accounted for in the analysis. The use of a three-level model highlights the potential to uncover mechanisms that bridge predictors and outcomes that reside at hierarchically distal levels of analysis in a theoretically sound manner.

Table 10. Two-Level Model Predicting Internal and External Team Process

	Internal team process		**External team process**	
Variable	1a	2a	1b	2b
Level-1:				
Intercept (β_{0j})	.16 (.055)	.14 (.064)	.07 (.019)	.05 (.025)
Project size (β_{1j})	.15** (.011)	.12* (.012)	.16** (.011)	.13* (.012)
Project complexity (β_{2j})	.13* (.014)	.10 (.035)	.18** (.019)	.13* (.022)
Requirements uncertainty (β_{3j})	.07 (.150)	.07 (.019)	.05 (.015)	.04 (.016)
Level-2:				
Transformational leadership (γ_{01})		.19*** (.010)		.20*** (.011)
Random effects:				
Level-1 variance (r_{ij})	.53***	.49***	.57***	.57***
Level-2 variance (U_{0j})	.40**	.33***	.43***	.38***
Deviance	6888.63	5617.75	4418.60	3975.83
χ^2	528.68	417.68	586.52	493.75
R^2	.06	.12	.06	.13

Notes:
1. Level-1, n=2,302; Level-2, n=234; Level-3, n=16.
2. * p < .05; ** p < .01; *** p < .001.

5.4.4 Application of Guideline 2.5

We have thus far illustrated how studies can represent relationships that cross three levels and that are hierarchically nested. We present here two potential variations of our illustration that call for cross-classified models, as discussed in Guideline 2.4.

In the first variation, if a developer worked on multiple projects for a specific software module and for all clients that purchased these modules, this creates a cross-classified context. Specifically, a software developer is identified by both clients and project teams. In such a situation, overlooking the multiple group membership would lead to an underspecified model that understates the group effect (Luo & Kwok, 2012). Researchers can use multiple membership models to more precisely conceptualize and model the nature of the context that involves plural association of lower-level entities with groups at higher levels.

As a second variation, consider a researcher who examines how team empowerment affects team process and job satisfaction of individual team members. If client liaisons rotate across clients during the study period (e.g., the vendor firm wants all their liaisons to develop rich experiences with all clients), this may create a changing context (or group effect) with respect to project teams, that needs to be considered and can be incorporated through the specification of a dynamic group model.

6 Limitations

Despite the significant potential of three-level modeling for IS research that has been discussed and illustrated so far, we should also be mindful of the limitations. First, three-level modeling requires significantly greater effort in theorizing. Research contexts may further add to the complexity to these issues. In addition to the complex process of partitioning variables into different levels, the researcher needs to decide what random effects and fixed effects should be included in a three-level model. Models with overly complicated random effects may not converge. Researchers are thus encouraged to study our guidelines and other resources to choose the best model based on theory, context, and model comparison techniques in order to develop the most meaningful and parsimonious model. Second, enough clusters and enough units in the clusters are needed to successfully identify a model and obtain unbiased estimates. For clustered data with very few clusters (i.e., when less than 15), fixed effect models may be preferred (McNeish & Stapleton, 2016). Third, the interpretation of three-level models requires careful thinking, especially for causal inferences from observational data (Gelman, 2006). However, this limitation can potentially be addressed by utilizing longitudinal data (Bingenheimer & Raudenbush, 2004) and consulting our discussion on endogeneity issues (see Guideline 2.6). Additionally, different centering techniques may affect the interpretation of the results (see Guideline 2.7) and thus appropriate judgment and caution are essential.

7 Conclusion

We highlighted the opportunity for IS researchers to engage more actively with context by incorporating contextual variables in their multilevel research through the use of three-level models. We also underscored the need for theory development and testing that bridges factors at hierarchically distal levels of analysis in an effort to understand the multiple levels at which IS phenomena unfold. In this research, we discussed the utility of three-level RCM for achieving these theoretically important ends and provided guidelines for linking theoretical objectives with specific multilevel analytical approaches.

The guidelines outlined in this article provide IS researchers a roadmap for explicitly incorporating elements of context into multilevel theory development and testing. Hierarchical context represents an important aspect of the IS phenomena we study. Ignoring the hierarchically nested structure of data in our research can present serious threats to the validity of empirical findings. Consideration of these guidelines ensures that IS researchers apply the appropriate degree of rigor to their empirical studies even when multilevel theorizing is not of interest. For those interested in taking the charge of explicitly theorizing about context, the guidelines emphasize the importance of using theory as a guide to determine if, and when, three-level models are appropriate for incorporating the influence of context. We noted several advantages that three-level models have over two-level models and alternative OLS models and outlined how IS researchers should combine theoretical consideration of context with research design and the hierarchical structure of the data when testing three-level models. We also offered guidance and an illustration of how cross-level moderation analysis could be used to test the influence of discrete elements of context on lower-level relationships. Overall, these guidelines offer a useful tool for researchers wishing to advance IS theories through the incorporation of discrete context.

Beyond considerations of context, the guidelines outlined in this essay also provide a means for researchers to bridge relationships between constructs that reside at hierarchically distal levels of analysis. We underscored the perils of drawing such linkages while skipping mechanisms at intervening levels of analysis and highlighted the value of three-level models, compared to two-level models, in enabling researchers to model the effects of mechanisms at such intervening levels. We also outlined procedures for testing such cross-level mediation. The approach outlined here ensures that IS researchers employ greater rigor in testing of cross-level moderation and cross-level mediation relationships.

Our guidelines hold great promise for future research in that they are applicable across a broad range of IS domains. It is well recognized that the IS discipline encompasses a diversity of technology-related topics and domains (Benbasat & Weber, 1996; Robey, 1996). This breadth in

IS topics also spans levels of analysis, with IS researchers examining phenomena at the country, industry, strategic group, community, business unit, team, and individual levels (Agarwal & Lucas, 2005). As the focus shifts between these different levels of analysis, naturally one researcher's micro-focus represents another researcher's macro-focus. For instance, a focus on online community design and long-term viability might constitute a micro-level focus for a macro-level researcher who is interested in country-level IT infrastructure. However, the same focus on online community design would constitute a macro-level focus for the individual-level researcher who is interested in the cognitions underlying contribution behavior in such communities. As another example, e-government could be studied at the macro-level to examine how such systems impact corruption control (Srivastava et al., 2016), social divide (Srivastava & Teo, 2007), and business competitiveness (Srivastava & Teo, 2007, 2008). It could also be studied at the micro-level to examine how citizens develop trust on e-government websites (Srivastava & Teo, 2009), which may in turn influence user perception about and use intention of these websites (Nishant et al., 2019; Teo et al., 2008). The guidelines outlined in this essay can be applied by both researchers, to the degree that they are interested in incorporating multiple levels of analysis into their investigation. The guidelines outlined here also provide the opportunity for IS researchers who work at different levels of analysis to collaborate and build bridges across various levels. The field as a whole stands to gain tremendously from ideas that might spring from such collaborative efforts (Straub, 2009).

In sum, although prior literature has utilized three-level models for hypothesis testing, there has been little to no discussion on the theoretical grounds surrounding the use of such models. In this research essay, we show how to connect theoretical objectives with the use of three-level models to (1) incorporate discrete context into multilevel theory development and testing and (2) uncover mechanisms linking factors at hierarchically distal levels of analysis. Recently, much has been made of the need to bridge the micro-macro gap in organizational research (Bamberger, 2008; Hitt et al., 2007; Mathieu & Chen, 2011; Rousseau, 2011). The guidelines outlined here offer an important way for researchers to bridge this gap and model the hierarchically structured complexity that is reflected in the social phenomena that IS researchers study. Beyond hierarchically structured complexity, three-level models can be employed to represent contexts where cross-classification of entities manifests as multiple membership in groups or dynamic membership of groups. The ubiquity of IT at multiple levels of analysis (i.e., individual, team, organization, community, industry, country) makes IS phenomena particularly ripe for taking the lead in developing and testing multilevel theories, thereby bridging the micro-macro gap.

7 References

Agarwal, R., & Lucas, H. C. (2005). The information systems identity crisis: Focusing on high-visibility and high-impact research. *MIS Quarterly*, *29*(3), 381–398.

Ahuja, M. K., & Thatcher, J. B. (2005). Moving beyond intentions and toward the theory of trying: Effects of work environment and gender on post-adoption information technology use. *MIS Quarterly*, *29*(3), 427–459.

Algina, J., & Swaminathan, H. (2011). Centering in two-level nested designs. In J. Hox & K. Roberts (Eds.), *Handbook of advanced multilevel analysis* (pp. 285-312). New York: Routledge.

Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. *Academy of Management Review*, *36*(2), 247–271.

Ancona, D. G. (1990). Outward bound: Strategies for team survival in an organization. *Academy of Management Journal*, *33*(2), 334–365.

Ancona, D. G., & Caldwell, D. F. (1992). Bridging the boundary: External activity and performance in organizational teams. *Administrative Science Quarterly*, *37*(4), 634–665.

Ang, S., Slaughter, S., & Ng, K. Y. (2002). Human capital and institutional determinants of information technology compensation: Modeling multilevel and cross-level interactions. *Management Science*, *48*(11), 1427–1445.

Antonakis, J., Bastardoz, N., & Rönkkö, M. (2021). On ignoring the random effects assumption in multilevel models: Review, critique, and recommendations. *Organizational Research*

Methods, *24*(2), 443–483.

Bamberger, P. (2008). From the editors: Beyond contextualization: Using context theories to narrow the micro-macro gap in management research. *Academy of Management Journal*, *51*(5), 839–846.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Bauer, D. J., Gottfredson, N. C., Dean, D., & Zucker, R. A. (2013). Analyzing repeated measures data on individuals nested within groups: accounting for dynamic group effects. *Psychological Methods*, *18*(1), 1–14.

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods*, *11*(2), 142–163.

Bélanger, F., Cefaratti, M., Carte, T., & Markham, S. E. (2014). Multilevel research in information systems: Concepts, strategies, problems, and pitfalls. *Journal of the Association for Information Systems*, *15*(9), 614–650.

Benbasat, I., & Weber, R. (1996). Research commentary: Rethinking "diversity" in information systems research. *Information Systems Research*, *7*(4), 389–399.

Bingenheimer, J. B., & Raudenbush, S. W. (2004). Statistical and substantive inferences in public health: Issues in the application of multilevel models. *Annual Review of Public Health*, *25*, 53–77.

Bliese, P. D., & Hanges, P. J. (2004). Being both too liberal and too conservative: The perils of treating grouped data as though they were independent. *Organizational Research Methods*, *7*(4), 400–417.

Boh, W. F., Slaughter, S. A., & Espinosa, J. A. (2007). Learning from experience in software development: A multilevel analysis. *Management Science*, *53*(8), 1315–1331.

Brohman, K., Addas, S., Dixon, J., & Pinsonneault, A. (2020). Cascading feedback: A longitudinal study of a feedback ecosystem for telemonitoring patients with chronic disease. *MIS Quarterly*, *44*(1), 421–450.

Burton-Jones, A., & Gallivan, M. J. (2007). Toward a deeper understanding of system usage in organizations: A multilevel perspective. *MIS Quarterly*, *31*(4), 657–679.

Cafri, G., Hedeker, D., & Aarons, G. A. (2015). An introduction and integration of cross-classified, multiple membership, and dynamic group random-effects models. Psychological methods, 20(4), 407–421.

Cappelli, P., & Sherer, P. D. (1991). The missing role of context in OB: The need for a meso-level approach. *Research in Organizational Behavior*, *13*, 55–110.

Chen, G., Kirkman, B. L., Kanfer, R., Allen, D., & Rosen, B. (2007). A multilevel study of leadership, empowerment, and performance in teams. *Journal of Applied Psychology*, *92*(2), 331–346.

Chiasson, M. W., & Davidson, E. (2015). Taking industry seriously in information systems research. *MIS Quarterly*, *29*(4), 591–605.

Colazo, J., & Fang, Y. (2010). Following the sun: Temporal dispersion and performance in open source software project teams. *Journal of the Association for Information Systems*, *11*(11), 684–707.

Corley, K. G., & Gioia, D. A. (2011). Building theory about theory building: what constitutes a theoretical contribution. *Academy of Management Review*, *36*(1), 12–32.

Dansereau, F., Alutto, J. A., & Yammarino, F. J. (1984). *Theory testing in organizational behavior: The variant approach*. Prentice-Hall.

Davison, M. L., Kwak, N., Seo, Y. S., & Choi, J. (2002). Using hierarchical linear models to examine moderator effects: Person-by-organization interactions. *Organizational Research Methods*, *5*(3), 231–254.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods*, *12*(2), 121–138.

Fang, Y., & Neufeld, D. (2009). Understanding sustained participation in open source software projects. *Journal of Management Information Systems*, *25*(4), 9–50.

Faraj, S., & Sambamurthy, V. (2006). Leadership of information systems development projects. *IEEE Transactions on Engineering Management*, *53*(2), 238–249.

Faraj, S., & Sproull, L. (2000). Coordinating expertise in software development teams. *Management Science*, *46*(12), 1554–1568.

Fielding, A., & Goldstein, H. (2006). Cross-classified and multiple membership structures in multilevel models: An introduction and review. DfES, University of Birmingham, 1–66.

Gelman, A. (2006). Multilevel (hierarchical) modeling: What it can and cannot do. *Technometrics*, *48*(3), 432–435.

Gopal, A., & Gosain, S. (2010). Research note: The role of organizational controls and boundary spanning in software development outsourcing: Implications for project performance. *Information Systems Research*, *21*(4), 960–982.

Gregor, S. (2006). The nature of theory in information systems. *MIS Quarterly*, *30*(3), 611–642.

Griffin, M. A. (2007). Specifying organizational contexts: systematic links between contexts and processes in organizational behavior. *Journal of Organizational Behavior*, *28*(7), 859–863.

Hackman, J. R. (2003). Learning more by crossing levels: Evidence from airplanes, hospitals, and orchestras. *Journal of Organizational Behavior*, *24*(8), 905–922.

Hackman, J. R., & Wageman, R. (2005). A theory of team coaching. *Academy of Management Review*, *30*(2), 269–287.

Heck, R. H., & Thomas, S. L. (2020). *An introduction to multilevel modeling techniques* (4th ed.). Routledge.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*(1), 153–161.

Herbsleb, J. D., & Mockus, A. (2003). An empirical study of speed and communication in globally distributed software development. *IEEE Transactions on Software Engineering*, *29*(6), 481–494.

Hitt, M. A., Beamish, P. W., Jackson, S. E., & Mathieu, J. E. (2007). Building theoretical and empirical bridges across levels: Multilevel research in management. *Academy of Management Journal*, *50*(6), 1385–1399.

Hong, W., Chan, F. K. Y., Thong, J. Y. L., Chasalow, L. C., & Dhillon, G. (2014). A framework and guidelines for context-specific theorizing in information systems research. *Information Systems Research*, *25*(1), 111–136.

House, R., Rousseau, D. M., & Thomas-Hunt, M. (1995). The meso paradigm: A framework for the integration of micro and macro organizational behavior. *Research in Organizational Behavior*, *17*, 71–114.

James, L. R., & Williams, L. J. (2000). The cross-level operator in regression, ANCOVA, and contextual analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 382–424). Jossey-Bass.

Jiang, Y., & Chen, C. C. (2018). Integrating knowledge activities for team innovation: Effects of transformational leadership. *Journal of Management*, *44*(5), 1819–1847.

Johns, G. (2001). In praise of context. *Journal of Organizational Behavior*, *22*(1), 31–42.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review*, *31*(2), 386–408.

Kane, G. C., & Borgatti, S. P. (2011). Centrality-IS proficiency alignment and workgroup performance. *MIS Quarterly*, *35*(4), 1063–1078.

Kane, G. C., & Labianca, G. (2011). IS avoidance in health-care groups: A multilevel investigation. *Information Systems Research*, *22*(3), 504–522.

Kang, S., Lim, K. H., Kim, M. S., & Yang, H.-D. (2012). Research note: A multilevel analysis of the effect of group appropriation on collaborative technologies use and performance. *Information Systems Research*, *23*(1), 214–230.

Kirsch, L. J., Ko, D.-G., & Haney, M. H. (2010). Investigating the antecedents of team-based clan control: Adding social capital as a predictor. *Organization Science*, *21*(2), 469–489.

Klein, K. J., Dansereau, F., & Hall, R. J. (1994). Levels issues in theory development, data collection, and analysis. *Academy of Management Review*, *19*(2), 195–229.

Klein, K. J., & Kozlowski, S. W. J. (2000). *Multilevel theory, research and methods in organizations*. Jossey-Bass.

Kohli, R., & Grover, V. (2008). Business value of IT: An essay on expanding research directions to keep up with the times. *Journal of the Association for Information Systems*, *9*(1), 23–39.

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). Jossey-Bass.

Kozlowski, S. W. J., & Salas, E. (1997). A multilevel organizational systems approach for the implementation and transfer of training. In J. K. Ford (Ed.), *Improving training effectiveness in work organizations* (pp. 247–287). Psychology Press.

Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research*, *36*(2), 249–277.

Lapointe, L., & Rivard, S. (2005). A multilevel model of resistance to information technology implementation. *MIS Quarterly*, *29*(3), 461–491.

Levina, N., & Vaast, E. (2008). Innovating or doing as told? Status differences and overlapping boundaries in offshore collaboration. *MIS Quarterly*, *32*(2), 307–332.

Luke, D. A. (2004). *Multilevel modeling*. Sage.

Luo, W., & Kwok, O.-M. (2012). The consequences of ignoring individuals' mobility in multilevel growth models: A Monte Carlo Study. Journal of Educational and Behavioral Statistics, 37(1), 31–56.

Ma, X., Kim, S. H., & Kim, S. S. (2014). Online gambling behavior: The impacts of cumulative outcomes, recent outcomes, and prior use. *Information Systems Research*, *25*(3), 511–527.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*(1), 83–104.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, *26*(3), 356–376.

Marrone, J. A., Tesluk, P. E., & Carson, J. B. (2007). A multilevel investigation of antecedents and consequences of team member boundary-spanning behavior. *Academy of Management Journal*, *50*(6), 1423–1439.

Maruping, L. M., & Magni, M. (2015). Motivating employees to explore collaboration technology in team contexts. *MIS Quarterly*, *39*(1), 1–16.

Maruping, L. M., Venkatesh, V., & Agarwal, R. (2009). A control theory perspective on agile methodology use and changing user requirements. *Information Systems Research*, *20*(3), 377–399.

Mathieu, J. E., & Chen, G. (2011). The etiology of the multilevel paradigm in management research. *Journal of Management*, *37*(2), 610–641.

Mathieu, J. E., & Taylor, S. R. (2007). A framework for testing meso-mediational relationships in organizational behavior. *Journal of Organizational Behavior*, *28*(2), 141–172.

McNeish, D., & Stapleton, L. M. (2016). Modeling clustered data with very few clusters. *Multivariate Behavioral Research*, *51*(4), 495–518.

Medappa, P. K., & Srivastava, S. C. (2019). Does superposition influence the success of FLOSS projects? An examination of open-source software development by organizations and individuals. *Information Systems Research*, *30*(3), 764–786.

Misangyi, V. F., Elms, H., Greckhamer, T., & Lepine, J. A. (2006). A new perspective on a fundamental debate: A multilevel approach to industry, corporate, and business unit effects.

Strategic Management Journal, *27*(6), 571–590.

Moerbeek, M. (2004). The consequence of ignoring a level of nesting in multilevel analysis. *Multivariate Behavioral Research*, *39*(1), 129–149.

Morgeson, F. P., & Hofmann, D. A. (1999). The structure and function of collective constructs: Implications for multilevel research and theory development. *Academy of Management Review*, *24*(2), 249–265.

Mowday, R. T., & Sutton, R. I. (1993). Organizational behavior: Linking individuals and groups to organizational contexts. *Annual Review of Psychology*, *44*(1), 195–229.

Nan, N. (2011). Capturing bottom-up information technology use processes: A complex adaptive systems model. *MIS Quarterly*, *35*(2), 505–532.

Nishant, R., Srivastava, S. C., & Teo, T. S. H. (2019). Using polynomial modeling to understand service quality in e–government websites. *MIS Quarterly*, *43*(3), 807–826.

Ou, A. Y., Seo, J., Choi, D., & Hom, P. W. (2017). When can humble top executives retain middle managers? The moderating role of top management team faultlines. *Academy of Management Journal*, *60*(5), 1915–1931.

Pituch, K. A., Murphy, D. L., & Tate, R. L. (2010). Three-level models for indirect effects in school- and class-randomized experiments in education. *Journal of Experimental Education* *78*(1), 60–95.

Preacher, K. J. (2011). Multilevel SEM strategies for evaluating mediation in three-level data. *Multivariate Behavioral Research*, *46*(4), 691–731.

Rabe-Hesketh, S., & Skrondal, A. (2021). *Multilevel and longitudinal modeling using Stata*. College Station, Texas: STATA Press.

Rai, A., Maruping, L. M., & Venkatesh, V. (2009). Offshore information systems project success: The role of social embeddedness and cultural characteristics. *MIS Quarterly*, *33*(3), 617–641.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Sage.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. T. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Scientific Software International.

Raykov, T. (2010). Proportion of third-level variation in multi-level studies: A note on an interval estimation procedure. *British Journal of Mathematical and Statistical Psychology*, *63*(2), 417–426.

Roberts, K. H., Hulin, C. L., & Rousseau, D. M. (1978). *Developing an interdisciplinary science of organizations*. Jossey-Bass.

Robey, D. (1996). Research Commentary: Diversity in information systems research: Threat, promise, and responsibility. *Information Systems Research*, *7*(4), 400–408.

Rousseau, D. M. (1985). Issues of level in organizational research: Multi-level and cross-level perspectives. *Research in Organizational Behavior*, *7*, 1–37.

Rousseau, D. M. (2011). Reinforcing the micro/macro bridge: Organizational thinking and pluralistic vehicles. *Journal of Management*, *37*(2), 429–442.

Rousseau, D. M., & Fried, Y. (2001). Location, location, location: Contextualizing organizational research. *Journal of Organizational Behavior*, *22*(1), 1–13.

Sasidharan, S., Santhanam, R., Brass, D. J., & Sambamurthy, V. (2012). The effects of social network structure on enterprise systems success: A longitudinal multilevel analysis. *Information Systems Research*, *23*(3), 658–678.

Seibert, S. E., Silver, S. R., & Randolph, W. A. (2004). Taking empowerment to the next level: A multiple-level model of empowerment, performance, and satisfaction. *Academy of Management Journal*, *47*(3), 332–349.

Shi, Y., Leite, W., & Algina, J. (2010). The impact of omitting the interaction between crossed factors in cross-classified random effects modelling. British Journal of Mathematical and Statistical Psychology, 63(1), 1–15.

Snijders, T. A. B., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Sage.

Srivastava, S. C., & Teo, T. S. H. (2007). E-government payoffs: Evidence from cross-country data. *Journal of Global Information Management*, *15*(4), 20–40.

Srivastava, S. C., & Teo, T. S. H. (2008). The relationship between e-government and national competitiveness: The moderating influence of environmental factors. *Communications of the Association for Information Systems*, *23*(1), 73–94.

Srivastava, S. C., & Teo, T. S. H. (2009). Citizen trust development for e-government adoption and usage: Insights from young adults in Singapore. *Communications of the Association for Information Systems*, *25*(1), 359–378.

Srivastava, S. C., & Teo, T. S. H. (2010). E-government, e-business, and national economic performance. *Communications of the Association for Information Systems*, *26*(1), 267–286.

Srivastava, S. C., Teo, T. S. H., & Devaraj, S. (2016). You can’t bribe a computer: Dealing with the societal challenge of corruption through ICT. *MIS Quarterly*, *40*(2), 511–526.

Straub, D. W. (2009). Editor’s comments: Creating blue oceans of thought via highly citable articles. *MIS Quarterly*, *36*(4), iii–vii.

Tarasewich, P., & Warkentin, M. (2002). Information everywhere. *Information Systems Management*, *19*(1), 8–13.

Teo, T. S. H., Srivastava, S. C., & Jiang, L. (2008). Trust and electronic government success: An empirical study. *Journal of Management Information Systems*, *25*(3), 99–132.

Thatcher, J. B., Brower, R. S., & Mason, R. M. (2006). Organizational fields and the diffusion of information technologies within and across the nonprofit and public sectors: A preliminary theory. *American Review of Public Administration*, *36*(4), 437–454.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478.

Venkatesh, V., Rai, A., & Maruping, L. M. (2018). Information systems projects and individual developer outcomes: Role of project managers and process control. *Information Systems*

Research, *29*(1), 127–148.

Venkatesh, V., Sykes, T. A., & Zhang, X. (2020). ICT for development in rural India: A longitudinal study of women's health outcomes. *MIS Quarterly*, *44*(2), 605–629.

Walz, D. B., Elam, J. J., & Curtis, B. (1993). Inside a software design team: Knowledge acquisition, sharing, and integration. *Communications of the ACM*, *36*(10), 63–77.

Wang, J., Shan, Z., Gupta, M., & Rao, H. R. (2019). A longitudinal study of unauthorized access attempts on information systems: The role of opportunity contexts. *MIS Quarterly*, *43*(2), 601–622.

Warkentin, M., Moore, R. S., Bekkering, E., & Johnston, A. C. (2009). Analysis of systems development project risks: An integrative framework. *DATA BASE for Advances in Information Systems*, *40*(2), 8–27.

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, *21*(1), 1–19.

Westner, M., & Strahringer, S. (2010). Determinants of success in IS offshoring projects: Results from an empirical study of German companies. *Information & Management*, *47*(5), 291–299.

Whetten, D. A. (2002). Modeling as theorizing: A systematic methodology for theory development. In D. Partington (Ed.), *Essential skills for management research* (pp. 45–71). Thousand Oaks, CA: Sage Publications.

Windeler, J. B., Maruping, L., & Venkatesh, V. (2017). Technical systems development risk factors: The role of empowering leadership in lowering developers' stress. *Information Systems Research*, *28*(4), 775–796.

Xie, K., & Lee, Y.-J. (2015). Social media and brand purchase: Quantifying the effects of exposures to earned and owned social media activities in a two-stage decision making model. *Journal of Management Information Systems*, *32*(2), 204–238.

Yau, L., Lee, S.-Y., & Poon, W.-Y. (1993). Covariance structure analysis with three level data. *Computational Statistics & Data Analysis*, *15*(2), 159–178.

Zhang, M., & Gable, G. G. (2017). A systematic framework for multilevel theorizing in information systems research. *Information Systems Research*, *28*(2), 203–224.

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods*, *12*(4), 695–719.

Zhao, K., Zhang, B., & Bai, X. (2018). Estimating contextual motivating factors in virtual interorganizational communities of practice: Peer effects and organizational influences. *Information Systems Research*, *29*(4), 910–927.

APPENDIX A

The sections below outline the equations used to (a) determine where the variance resides, (b) test for cross-level moderation, and (c) test for cross-level mediation in three-level RCM. Note that although only the equations for developer performance are shown, the same equations would be used in the case of psychological stress as well.

A. Equations Used in Determining Where the Variance Resides

The following formulation of an unconditional model at the first level (individual-level) would help accomplish this:

$$Y_{ijk} = \pi_{0jk} + e_{ijk} \text{ (1a) Level-1}$$

where Y_{ijk} represents the individual performance for developer i (i = 1, 2,…n_{jk}) who works in project team j (j = 1, 2,…J_k) under client liaison k (k = 1, 2,…K), π_{0ij} represents the mean level of developer performance within project team j under client liaison k, and e_{ijk} represents the random deviation of developer *ijk*'s performance score from the mean level of performance among developers in their project team. The error is assumed to be normally distributed with a mean of 0 and variance σ^2.

At the second level, the mean level of performance within each project team (π_{0jk}), is modeled as an outcome that can vary randomly around a client liaison level mean, using the following formulation:

$$\pi_{0jk} = \beta_{00k} + r_{0jk} \text{ (1b) Level-2}$$

where β_{00k} represents the mean level of developer performance under client liaison k and r_{0jk} represents the deviation of mean developer performance in project team *jk* from the mean level of developer performance at the client liaison level. Here too, the error is assumed to be normally distributed with mean 0 and variance τ_π. The variability in developer performance in project teams under each of the *k* client liaisons is assumed to be the same (Raudenbush & Bryk, 2002).

Finally, the variability in developer performance among client liaisons can be modeled by allowing the client liaison level means (β_{00k}) to vary randomly around a grand mean using the following formulation:

$$\beta_{00k} = \gamma_{000} + u_{00k} \text{ (1c) Level-3}$$

where γ_{000} represents the grand mean and u_{00k} represents the deviation of client liaison *k*'s mean from the grand mean. As with the other two levels, the error is assumed to be normally distributed with mean 0 and variance τ_β.

This three-level model allows the researcher to decompose the variability in developer performance into its components across all three levels of analysis. Using the following formulations, the researcher can determine how much of the variability is accounted for by each level of analysis:

$$\sigma^2 / (\sigma^2 + \tau_\pi + \tau_\beta) \text{ (2a) Level-1}$$

which represents the proportion of variance at the individual level (or the within-project team variance);

$$\tau_\pi / (\sigma^2 + \tau_\pi + \tau_\beta) \text{ (2b) Level-2}$$

which represents the proportion of variance at the project team level (or the between-project team variance within client liaisons); and

$$\tau_\beta / (\sigma^2 + \tau_\pi + \tau_\beta) \text{ (2c) Level-3}$$

which represents the proportion of variance at the client liaison level.

B. Equations Used in Testing for Cross-level Moderation (Level-2 X Level-3)

Table 11: Models Predicting Individual Developer Performance

MODEL 1
Level-1:
Developer performance$_{ijk}$ = $\pi_{0jk} + \pi_{1jk}$ Conscientiousness + π_{2jk} Extraversion + e_{ijk}

Level-2:
$\pi_{0jk} = \beta_{00k} + \beta_{01k}$ Project size + β_{02k} Project complexity + β_{03k} Requirements uncertainty + r_{0jk};
$\pi_{1jk} = \beta_{10k}$; $\pi_{2jk} = \beta_{20k}$

Level-3:
$\beta_{00k} = \gamma_{000} + U_{00k}$; $\beta_{01k} = \gamma_{010}$; $\beta_{02k} = \gamma_{020}$; $\beta_{03k} = \gamma_{030}$

MODEL 2 (With Coefficients for Team Processes Treated as Fixed)
Level-1:
Developer performance$_{ijk}$ = $\pi_{0jk} + \pi_{1jk}$ Conscientiousness + π_{2jk} Extraversion + e_{ijk}

Level-2:
$\pi_{0jk} = \beta_{00k} + \beta_{01k}$ Project size + β_{02k} Project complexity + β_{03k} Requirements uncertainty + β_{04k} Internal team process + β_{05k} External team process + r_{0jk}; $\pi_{1jk} = \beta_{10k}$; $\pi_{2jk} = \beta_{20k}$

Level-3:
$\beta_{00k} = \gamma_{000} + U_{00k}$; $\beta_{01k} = \gamma_{010}$; $\beta_{02k} = \gamma_{020}$; $\beta_{03k} = \gamma_{030}$; $\beta_{04k} = \gamma_{040}$; $\beta_{05k} = \gamma_{050}$

MODEL 3 (With Coefficients for Team Processes Treated as Random)
Level-1:
Developer performance$_{ijk}$ = $\pi_{0jk} + \pi_{1jk}$ Conscientiousness + π_{2jk} Extraversion + e_{ijk}

Level-2:
$\pi_{0jk} = \beta_{00k} + \beta_{01k}$ Project size + β_{02k} Project complexity + β_{03k} Requirements uncertainty + β_{04k} Internal team process + β_{05k} External team process + r_{0jk}; $\pi_{1jk} = \beta_{10k}$; $\pi_{2jk} = \beta_{20k}$

Level-3:
$\beta_{00k} = \gamma_{000} + U_{00k}$; $\beta_{01k} = \gamma_{010}$; $\beta_{02k} = \gamma_{020}$; $\beta_{03k} = \gamma_{030}$; $\beta_{04k} = \gamma_{040} + U_{04k}$; $\beta_{05k} = \gamma_{050} + U_{05k}$

MODEL 4
Level-1:
Developer performance$_{ijk}$ = $\pi_{0jk} + \pi_{1jk}$ Conscientiousness + π_{2jk} Extraversion + e_{ijk}

Level-2:
$\pi_{0jk} = \beta_{00k} + \beta_{01k}$ Project size + β_{02k} Project complexity + β_{03k} Requirements uncertainty + β_{04k} Internal team process + β_{05k} External team process + r_{0jk}; $\pi_{1jk} = \beta_{10k}$; $\pi_{2jk} = \beta_{20k}$

Level-3:
$\beta_{00k} = \gamma_{000} + \gamma_{001}$ Client liaison experience + U_{00k}; $\beta_{01k} = \gamma_{010}$; $\beta_{02k} = \gamma_{020}$; $\beta_{03k} = \gamma_{030}$; $\beta_{04k} = \gamma_{040} + \gamma_{041}$ Client liaison experience + U_{04k}; $\beta_{05k} = \gamma_{050} + \gamma_{051}$ Client liaison experience + U_{05k}

C. Equations Used in Testing for Cross-level Mediation

Table 12: Models Predicting Individual Developer Performance (Level-3/Level-2→Level-1)

MODEL 1

Level-1:

Developer performance$_{ijk}$ = π_{0jk} + π_{1jk} Conscientiousness + π_{2jk} Extraversion + e_{ijk}

Level-2:

$\pi_{0jk} = \beta_{00k} + \beta_{01k}$ Project size + β_{02k} Project complexity + β_{03k} Requirements uncertainty + r_{0jk};
$\pi_{1jk} = \beta_{10k}$; $\pi_{2jk} = \beta_{20k}$

Level-3:

$\beta_{00k} = \gamma_{000} + U_{00k}$; $\beta_{01k} = \gamma_{010}$; $\beta_{02k} = \gamma_{020}$; $\beta_{03k} = \gamma_{030}$

MODEL 2

Level-1:

Developer performance$_{ijk}$ = π_{0jk} + π_{1jk} Conscientiousness + π_{2jk} Extraversion + e_{ijk}

Level-2:

$\pi_{0jk} = \beta_{00k} + \beta_{01k}$ Project size + β_{02k} Project complexity + β_{03k} Requirements uncertainty + r_{0jk};
$\pi_{1jk} = \beta_{10k}$; $\pi_{2jk} = \beta_{20k}$

Level-3:

$\beta_{00k} = \gamma_{000} + \gamma_{001}$ Transformational leadership + U_{00k}; $\beta_{01k} = \gamma_{010}$; $\beta_{02k} = \gamma_{020}$; $\beta_{03k} = \gamma_{030}$

MODEL 3

Level-1:

Developer performance$_{ijk}$ = π_{0jk} + π_{1jk} Conscientiousness + π_{2jk} Extraversion + e_{ijk}

Level-2:

$\pi_{0jk} = \beta_{00k} + \beta_{01k}$ Project size + β_{02k} Project complexity + β_{03k} Requirements uncertainty + β_{04k} Internal team process + β_{05k} External team process + r_{0jk}; $\pi_{1jk} = \beta_{10k}$; $\pi_{2jk} = \beta_{20k}$

Level-3:

$\beta_{00k} = \gamma_{000} + U_{00k}$; $\beta_{01k} = \gamma_{010}$; $\beta_{02k} = \gamma_{020}$; $\beta_{03k} = \gamma_{030}$; $\beta_{04k} = \gamma_{040}$; $\beta_{05k} = \gamma_{050}$

MODEL 4

Level-1:

Developer performance$_{ijk}$ = π_{0jk} + π_{1jk} Conscientiousness + π_{2jk} Extraversion + e_{ijk}

Level-2:

$\pi_{0jk} = \beta_{00k} + \beta_{01k}$ Project size + β_{02k} Project complexity + β_{03k} Requirements uncertainty + β_{04k} Internal team process + β_{05k} External team process + r_{0jk}; $\pi_{1jk} = \beta_{10k}$; $\pi_{2jk} = \beta_{20k}$

Level-3:

$\beta_{00k} = \gamma_{000} + \gamma_{001}$ Transformational leadership + U_{00k}; $\beta_{01k} = \gamma_{010}$; $\beta_{02k} = \gamma_{020}$; $\beta_{03k} = \gamma_{030}$; $\beta_{04k} = \gamma_{040}$; $\beta_{05k} = \gamma_{050}$

Table 13: Models Predicting Team Processes (Level-3 → Level-2)

(Note: Equations are the same for models predicting external team process)

MODEL 1
Level-1:
Internal team process$_{ij}$ = β_{0j} + β_{1j} Project size + β_{2j} Project complexity + β_{3j} Requirements uncertainty + r_{ij}

Level-2:
$\beta_{0j} = \gamma_{00} + U_{0j}$; $\beta_{1j} = \gamma_{10}$; $\beta_{2j} = \gamma_{20}$; $\beta_{3j} = \gamma_{30}$

MODEL 2
Level-1:
Internal team process$_{ij}$ = β_{0j} + β_{1j} Project size + β_{2j} Project complexity + β_{3j} Requirements uncertainty + r_{ij}

Level-2:
$\beta_{0j} = \gamma_{00} + \gamma_{01}$ Transformational leadership $+ U_{0j}$; $\beta_{1j} = \gamma_{10}$; $\beta_{2j} = \gamma_{20}$; $\beta_{3j} = \gamma_{30}$

References

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Sage.

APPENDIX B

Measurement

Dependent Variables

Individual performance ratings were provided by the project leader for every individual team member and were based on a proprietary 4-item scale used in the organization. The measure was designed to evaluate various aspects of a team member's work on a particular project and included items, such as quality of work output, quantity of work output and effort put forth on the project. The vendor firm used these ratings on this scale in its performance appraisals. The scale had a reliability of .75.

Psychological stress was measured using a scale adapted from Keller (2001). The scale consists of four items that capture the extent to which an individual experiences an aversive psychological response to their work environment. We adapted the scale to measure psychological stress relating to the focal project. Team members provided responses to the psychological stress items at the end of the project. The reliability of the scale was .71.

Independent Variables

The *transformational leadership* of client liaisons was measured on a 15-point scale that has multiitem subscales corresponding to five transformational leadership behaviors. The scale captures the extent to which a client liaison adopts leadership behaviors including leading by example, participative decision making, coaching, informing, and showing concern (Srivastava et al., 2006). Higher scores of transformational leadership denote greater endorsement of empowering leadership values. *Client liaison experience* was measured by obtaining archival records on the number of projects a specific client liaison had managed prior to the current project.

Given our focus on the extent to which project teams engaged in interactions with their clients, we measured *external team processes* using a five-item subscale of Ancona and Caldwell's (1992) boundary spanning scale. Specifically, we utilized the ambassadorial activity subscale, which reflects the extent to which a team interacts with its external constituents to understand their needs and obtain their feedback. The scale was adapted to fit the context of the study by referring to the team's interactions with the client firm. The scale demonstrated an adequate level of reliability (α = .75). Because multiple team members within each team responded to questions about external team processes, it was necessary to determine the extent to which these responses converged within teams. The average $r_{wg(j)}$ for external team processes exceeded the recommended cutoff of .70, suggesting that individual team member responses to this scale could be aggregated to compute a single team-level score (James et al., 1984). We computed a team-level score for external team processes by averaging the responses provided by members of a team. From project archives, we were also able to obtain an objective measure of the number of visits the project team made to the client firm over the course of the entire project. The correlation between this objective measure and the aggregated measure of external team process was high, giving us confidence that responses to the survey measure were reflective of the teams' actual activities.

Internal team processes were measured using a nine-item scale by Mathieu et al. (2006). The scale includes three dimensions—each consisting of three items—from Marks et al.'s (2001) super-ordinate categories of team processes: transition, action, and interpersonal processes. Collectively, the scale items capture the extent to which a team has established procedures for outlining core objectives, monitoring progress toward those objectives, tracking of resources, and management of team member well-being (Marks et al., 2001). The reliability of the scale was

.73. Results of a CFA indicated a good fit for the three first-order factors (transition, action, and interpersonal processes) and a single second-order factor (internal team processes). The average $r_{wg(j)}$ of over .70 suggested that it was appropriate to aggregate team member responses to represent a team-level score (James et al., 1984). Hence, we averaged team members' responses within each project team to compute a team-level score for internal team process.

Control Variables

Conscientiousness and *extraversion* are two individual differences that have been consistently linked to job performance (Barrick et al., 1998; Barrick & Mount, 1991). Therefore, we captured these individual differences using scales from Costa and McCrae's (1992) NEO-FFI personality scale, which is among the most widely used instruments for measuring the five-factor model. Each scale consists of 12 items. The reliabilities for conscientiousness and extraversion were .73 and .75 respectively.

Project characteristics have been found to be predictors of individual and team performance. We controlled for three structural aspects of the project: project size, project complexity, and requirements uncertainty, which have consistently been linked to project outcomes (see Keil et al., 2000; Maruping et al., 2009; Ravichandran & Rai, 2000; Wallace et al., 2004; Warkentin et al., 2009). As part of its standard routine, the vendor firm kept detailed records on various software project metrics for completed projects. We obtained project size, requirements uncertainty, and project complexity from the vendor firm's archival records. Consistent with prior research, *project size* was measured as total lines of code and *project complexity* was measured as the total number of adjusted function points, based on fourteen complexity characteristics that account for the different kinds of system requirements and development environments (see Albrecht & Gaffney Jr., 1983; Banker et al., 1998;

Mukhopadhyay et al., 1992). *Requirements uncertainty* was measured as the number of formal written changes to the original requirements.

References

Albrecht, A. J., & Gaffney Jr., J. E. (1983). Software function, source lines of code, and development effort prediction: A software science validation. *IEEE Transactions on Software Engineering*, *SE-9*(6), 639–648.

Ancona, D. G., & Caldwell, D. F. (1992). Bridging the boundary: External activity and performance in organizational teams. *Administrative Science Quarterly*, *37*(4), 634–665.

Banker, R. D., Davis, G. B., & Slaughter, S. A. (1998). Software development practices, software complexity, and software maintenance performance: A field study. *Management Science*, *44*(4), 433–450.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*(1), 1–26.

Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology*, *83*(3), 377–391.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, *69*(1), 85–98.

Keil, M., Mann, J., & Rai, A. (2000). Why software projects escalate: An empirical analysis and test of four theoretical models. *MIS Quarterly*, *24*(4), 631–664.

Keller, R. T. (2001). Cross-functional project groups in research and new product development: Diversity, communications, job stress, and outcomes. *Academy of Management Journal*, *44*(3), 547–555.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, *26*(3), 356–376.

Maruping, L. M., Venkatesh, V., & Agarwal, R. (2009). A control theory perspective on agile methodology use and changing user requirements. *Information Systems Research*, *20*(3), 377–399.

Mathieu, J. E., Gilson, L. L., & Ruddy, T. M. (2006). Empowerment and team effectiveness: An empirical test of an integrated model. *Journal of Applied Psychology*, *91*(1), 97–108.

Mukhopadhyay, T., Vicinanza, S. S., & Prietula, M. J. (1992). Examining the feasibility of a case-based reasoning model for software effort estimation. *MIS Quarterly*, *16*(2), 155–170.

Ravichandran, T., & Rai, A. (2000). Quality management in systems development: An organizational system perspective. *MIS Quarterly*, *24*(3), 381–415.

Srivastava, A., Bartol, K. M., & Locke, E. A. (2006). Empowering leadership in management teams: Effects on knowledge sharing, efficacy, and performance. *Academy of Management Journal*, *49*(6), 1239–1251.

Wallace, L., Keil, M., & Rai, A. (2004). How software project risk affects project performance: An investigation of the dimensions of risk and an exploratory model. *Decision Sciences*, *35*(2), 289–321.

Warkentin, M., Moore, R. S., Bekkering, E., & Johnston, A. C. (2009). Analysis of systems development project risks: An integrative framework. *DATA BASE for Advances in Information Systems*, *40*(2), 8–27.

APPENDIX C

Bélanger and colleagues (2014) conducted a thorough review of 526 papers published in two leading IS journals from 2002-2010 and found less than 10% used either multilevel theorizing or a multilevel research design. Without repeating their effort, we examined papers in the "Basket of Eight" journals between 2011 and 2020, including *MIS Quarterly*, *Information Systems Research*, *Journal of Management Information Systems,* the *Journal of the AIS*, *European Journal of Information Systems*, *Information Systems Journal*, *Journal of Information Technology*, and *Journal of Strategic Information Systems*. We identified 33 quantitative multilevel papers, 10 of which employed a three-level analysis, as listed in Table 14.

Table 14: Three-Level Analysis Papers in the "Basket of Eight" Journals (2011-2020)

Reference	Levels	Focal Topic	Data Type/ Study Design	Analysis Method	Level-3 Variables	Cross-level Interaction
Brohman et al. (2020)	Observation, individual, service	Healthcare feedback system	Archival; longitudinal	Generalized linear mixed model (GLMM)	Categorical	None
Kane and Borgatti (2011)	Patient, doctor, group	Alignment of centrality and IS proficiency	Survey; cross-sectional	Multiple regression corrected by Huber-White robust standard errors	Categorical & continuous	None
Kane and Labianca (2011)	Patient, doctor, group	IS avoidance	Survey; cross-sectional	Multiple regression corrected by Huber-White robust standard errors	Categorical & continuous	None
Ma et al. (2014)	Observation, individual, country	Online gambling behavior	Archival; longitudinal	HLM model	Categorical	None
Sasidharan et al. (2012)	Observation, individual, unit	Enterprise system success	Survey; longitudinal	HLM model	Continuous	Level-3 variable moderates level-1 relationship
Venkatesh et al. (2018)	Individual, project, manager	Project manager and process control	Survey; cross-sectional	HLM model	Continuous	Level-2 variable moderates level-2 to level-1 relationship
Venkatesh et al. (2020)	Observation, individual, village	ICT4D	Survey; longitudinal	HLM model	Continuous	Level-1 variable moderates level-2 to level-1 relationship & level-3 to level-1 relationship
Wang et al. (2019)	Observation, employee, department	System unauthorized access	Archival; longitudinal	Random-coefficient model with three levels	Continuous	Level-3 variable moderates level-1 relationships
Xie and Lee (2015)	Observation, household, brand	Exposure to social media on purchase behavior	Archival; longitudinal	Mixed effects model	Categorical & continuous	None
Zhao et al. (2018)	Individual, organization,	Virtual IOCoP community	Archival; longitudinal	Fixed effects structural model	Categorical & continuous	None

	aggregation	participation					
Note: All studies here are in top-down direction for three level modeling (i.e., dependent variables are at the lowest level).							

References

Bélanger, F., Cefaratti, M., Carte, T., & Markham, S. E. (2014). Multilevel research in information systems: Concepts, strategies, problems, and pitfalls. *Journal of the Association for Information Systems*, *15*(9), 614–650.

Brohman, K., Addas, S., Dixon, J., & Pinsonneault, A. (2020). Cascading feedback: A longitudinal study of a feedback ecosystem for telemonitoring patients with chronic disease. *MIS Quarterly*, *44*(1), 421–450.

Kane, G. C., & Borgatti, S. P. (2011). Centrality-IS proficiency alignment and workgroup performance. *MIS Quarterly*, *35*(4), 1063–1078.

Kane, G. C., & Labianca, G. (2011). IS avoidance in health-care groups: A multilevel investigation. *Information Systems Research*, *22*(3), 504–522.

Ma, X., Kim, S. H., & Kim, S. S. (2014). Online gambling behavior: The impacts of cumulative outcomes, recent outcomes, and prior use. *Information Systems Research*, *25*(3), 511–527.

Sasidharan, S., Santhanam, R., Brass, D. J., & Sambamurthy, V. (2012). The effects of social network structure on enterprise systems success: A longitudinal multilevel analysis. *Information Systems Research*, *23*(3-part-1), 658–678.

Venkatesh, V., Rai, A., & Maruping, L. M. (2018). Information systems projects and individual developer outcomes: Role of project managers and process control. *Information Systems Research*, *29*(1), 127–148.

Venkatesh, V., Sykes, T. A., & Zhang, X. (2020). ICT for development in rural India: A longitudinal study of women's health outcomes. *MIS Quarterly*, *44*(2), 605–629.

Wang, J., Shan, Z., Gupta, M., & Rao, H. R. (2019). A longitudinal study of unauthorized access attempts on information systems: The role of opportunity contexts. *MIS Quarterly*, *43*(2), 601–622.

Xie, K., & Lee, Y.-J. (2015). Social media and brand purchase: Quantifying the effects of exposures to earned and owned social media activities in a two-stage decision making model. *Journal of Management Information Systems*, *32*(2), 204–238.

Zhao, K., Zhang, B., & Bai, X. (2018). Estimating contextual motivating factors in virtual interorganizational communities of practice: Peer effects and organizational influences. *Information Systems Research*, *29*(4), 910–927.