



Looking at the bigger picture: on the relevance of null effects of model differences across time

Janna S. Ehrlich¹ · Charlotte Kreienbaum² · Christian M. Ringle^{2,3}

Received: 16 April 2026 / Accepted: 11 June 2026
© The Author(s) 2026

Abstract

Longitudinal research traditionally aims to detect and quantify change in outcomes such as attitudes, behaviors, or performance over time. Nonsignificant results, particularly in regards to longitudinal research designs, are however often overlooked, despite their potential theoretical and managerial relevance. This research contributes to addressing this gap by examining the stability of a theoretically grounded model across five independent time points. Adopting a model-logic approach to longitudinal analysis, time is conceptualized as a contextual dimension (e.g., a moderator) rather than as a predictor. Multigroup PLS-SEM is employed to assess the difference between the theoretically established relationships across years. The results do not reveal statistically significant differences in key structural paths, indicating a high degree of temporal stability. These null results provide evidence for the robustness and generalizability of the underlying theoretical mechanisms across varying economic and societal contexts. The study furthermore situates its design in the broader open science movement by framing longitudinal model comparisons as a form of structured temporal replication. Implications and guidelines for theory development, longitudinal modeling, and future research on replication and robustness in structural equation modeling are discussed.

Keywords Null effects · Nonsignificant effects · Longitudinal data · Longitudinal SEM · Open science

Introduction

Research in social sciences disciplines such as marketing and management often aims to reveal and statistically support significant effects, even though such effects are sometimes not particularly strong and relevant (e.g., Bagozzi and Dholakia 2002; Cannon and Homburg 2001; Melumad et al. 2019). This has led to the well-known “file drawer problem”, which laments that studies with significant results are

far more likely to be planned, conducted, and published than those with nonsignificant results (Franco et al. 2014; Rosenthal 1979). The file drawer problem is widely discussed in relation to the replication crisis (Open Science Collaboration 2015), as the selective publication of statistically significant findings can lead to an overrepresentation of false positive results in the published literature (Simmons et al. 2011).

As Petrescu and Krishen (2025) noted, a null result is not equivalent to “no result.” In a broad sense, it refers to an empirical finding that does not provide statistical support for a focal hypothesis. This typically corresponds to a nonsignificant effect or difference (i.e., a null effect) within the framework of null hypothesis significance testing (NHST; Gigerenzer et al. 2004; Wasserstein et al. 2019). Importantly, this should be distinguished from a null hypothesis in NHST, which conventionally specifies the absence of an effect as the default position against which an alternative hypothesis is evaluated (Wasserstein et al. 2019). By contrast, some research contexts involve a theoretically formulated stability or invariance hypothesis, in which the absence of differences across conditions or time is itself the expected

✉ Christian M. Ringle
c.ringle@tuhh.de

Janna S. Ehrlich
janna.ehrlich@hochschule-stralsund.de

Charlotte Kreienbaum
charlotte.kreienbaum@tuhh.de

¹ Hochschule Stralsund, Stralsund, Germany

² Institute of Management and Decision Sciences, Hamburg University of Technology, Am Schwarzenberg-Campus 4 (D), 21073 Hamburg, Germany

³ James Cook University, Townsville, Australia



outcome rather than a failure to reject a null hypothesis of no effect (Little 2013; Vandenberg and Lance 2000).

Not all null results have the same scientific meaning. When researchers expect change but observe nonsignificant effects, null results are only potentially informative if methodological requirements are met. In contrast, when theory explicitly predicts stability or invariance—as in longitudinal models of temporal robustness—nonsignificant differences can represent confirmatory evidence of the hypothesized pattern (e.g., Lakens 2017). Informative null results therefore refer to nonsignificant findings that are theoretically meaningful and methodologically credible, rather than to the formal testing of a null hypothesis in NHST (Landis et al. 2014). They arise when hypotheses are clearly specified *a priori*, measures and research designs are appropriate, statistical power is sufficient, and analyses are transparently reported. In this context, our research addresses both the interpretation of null effects in longitudinal structural models and the conditions under which such nonsignificant differences can be understood as evidence for temporal stability rather than inconclusive findings. Such findings can reduce research waste, refine theory by ruling out unsupported relationships, and inform practice when the absence of meaningful differences favors simpler or less costly alternatives, such as survey implementation choices (Kozlov 2024; Landis et al. 2014).

For example, when a more complex alternative model does not yield significantly better model fit or predictive power, this null difference supports the use of the more parsimonious theoretically established model (Liengaard et al. 2021; Satorra and Bentler 2010). Similarly, multigroup and moderator analyses (e.g., Cheah et al. 2023; Matthews 2017; Memon et al. 2019, 2026) may fail to provide evidence of observed heterogeneity. Likewise, latent class and segmentation analyses, such as FIMIX-PLS (Hahn et al. 2002; Hair et al. 2016; Matthews et al. 2016; Sarstedt et al. 2022b), may fail to reveal meaningful unobserved heterogeneity, for example, when no substantial heterogeneity or relevant segments can be identified. Such null findings support the model's general applicability across the studied population and strengthen confidence in the theoretically established model (Becker et al. 2013; Meredith 1993).

Longitudinal research designs are typically motivated by the expectation of change and are widely used to examine dynamic processes and developmental trajectories over time (Ployhart and Vandenberg 2010; Singer and Willett 2003; Bollen and Curran 2006). Comparatively less attention has been devoted to using longitudinal designs to assess the stability of theoretically specified relationships. Ployhart and Vandenberg (2010) defined longitudinal research as research designed to examine the dynamic nature of constructs by focusing on change rather than static representations of a

constructs. It involves collecting repeated measurements of the same variables from the same units of observation (e.g., individuals, teams, organizations) over time, with the ability to link the units across waves. This approach allows researchers to analyze both within-unit change (e.g., growth trajectories) and differences between units in how they change over time (Bollen and Curran 2006; Singer and Willett 2003). Typically, at least three time points are recommended to adequately capture these patterns of change (Chan 1998; Ployhart and Vandenberg 2010).

Our research contributes to the literature on informative null results by emphasizing their relevance in longitudinal research. In most empirical settings, null results arise when a study predicts an effect or difference but fails to detect one, thereby providing limited support for the original hypothesis. In longitudinal research, however, theories may predict stability rather than change. In such cases, the absence of significant differences across time is not a failure to confirm a hypothesized effect but is instead consistent with the hypothesized invariance of the underlying relationships. Prior research on ambient scent effects, for example, suggests that an initial change in consumers' perceptions of a service environment may persist over time and may not necessarily diminish even after the atmospheric cue has been removed (Girard et al. 2019; see also Biswas and Szocs 2019). Under these conditions, the theoretically relevant question is not whether relationships change, but whether they remain stable. We therefore distinguish between null results that emerge from an unsupported expectation of difference and null results that emerge from a supported expectation of stability. Consequently, longitudinal research should not focus exclusively on detecting change, but should also recognize the theoretical value of demonstrating temporal robustness. Accordingly, anticipating and empirically establishing the absence of significant differences in coefficients across time should be regarded as an important research objective—one that remains largely overlooked in current social science practice.

A further contribution of this research lies in its focus on longitudinal structural equation modeling (SEM), a central multivariate method in marketing (Hair et al. 2017; Sarstedt et al. 2022a). We introduce a classification that distinguishes between trend-logic and model-logic approaches to longitudinal SEM analysis. Focusing on the latter, we empirically examine whether structural model coefficients obtained by using partial least squares structural equation modeling (PLS-SEM; Wold 1982; Lohmöller 1989; Hair et al. 2027) remain stable over time, proposing a systematic analytical approach. Specifically, we test the invariance of an established model across five biennial waves, consistent with the theoretical expectation of stability. Based on this application, we derive practical guidance for longitudinal



PLS-SEM studies in which informative null differences are anticipated. Finally, we summarize key findings and outline promising avenues for future research on this important topic in marketing.

Brief overview of longitudinal research goals

Longitudinal research designs play a central role in social science research including the management and marketing disciplines, as organizational, market, and societal contexts evolve over time (Rindfleisch et al. 2008). They are traditionally employed to detect and quantify change, for example in attitudes, behaviors, or performance outcomes. Such approaches often focus on mean-level developments or growth trajectories and are particularly common in developmental and economic research (Singer and Willett 2003). Growth curve modeling and trend analyses are typically used when change itself constitutes the primary object of investigation (Bollen and Curran 2006). In contrast,

research in management and marketing often uses SEM to analyze theoretically established hypotheses for relationships among latent constructs (e.g., Diamantopoulos and Siguaw 2000; Rigdon 1998). These models are grounded in theoretical assumptions about causal mechanisms. Implicit in this approach is the assumption that the measurement and structural parameters are invariant across time, such that the theorized relationships are not specific to a single temporal context (Little 2013; Kline 2023). Empirical assessments of whether such structural relationships remain stable over time however remain comparatively rare.

The distinction between these longitudinal approaches reflects two fundamentally different analytical logics. A trend-logic approach assumes that systematic change is expected and seeks to model and quantify developmental trajectories over time (e.g., Collins 2006; Curran et al. 2010). In contrast, a model-logic approach evaluates whether theoretically specified measurement and structural parameters remain invariant or vary with time (e.g., Vandenberg and Lance 2000). In the latter approach, time does not function primarily as a predictor of outcomes but rather as a moderator against which structural stability is assessed (George and Jones 2000; see also Roemer 2016). Table 1 summarizes the differences.

The distinction builds on established longitudinal approaches that model time either as a predictor of change or as a grouping dimension for assessing invariance across time (e.g., Bollen and Curran 2006; Little 2013; Vandenberg and Lance 2000), with null findings interpreted as evidence of stability or invariance (e.g., Meehl 1990).

These two analytical logics are reflected in the different longitudinal modeling strategies available with SEM, offering several approaches for analyzing longitudinal data, depending on the theoretical objective (Kline 2023). For example, latent growth curve models are appropriate when the focus lies on intraindividual trajectories (Bollen and Curran 2006) and cross-lagged panel models examine reciprocal dynamics and temporal ordering (Hamaker et al. 2015). Both analysis types fall under the trend-logic approach, as they model time as predictors of the goal variable. Beyond single-study applications, meta-analytic structural equation modeling (meta-SEM) extends this logic by synthesizing such longitudinal relationships across studies (e.g., Nath et al. 2025).

When the objective is to assess whether the theoretically established relationships remain stable across time (e.g., using regression models, path models, or SEM), multigroup analyses (MGA) are particularly suitable because each measurement wave can be treated as a separate group and the estimated relationships can be compared across waves (Little 2013; Vandenberg and Lance 2000). Unlike growth models and trend analyses, which are primarily designed

Table 1 Trend-logic and model-logic longitudinal analysis comparison

Dimension	Trend-logic approach (time as predictor)	Model-logic approach (time as moderator)
Core aim of analysis	To test whether a focal construct systematically changes over time	To examine whether theoretically specified relationships remain stable or vary across time
Role of time	Time is an independent variable that predicts change in outcomes (e.g., linear or robust trends)	Time defines comparison groups; it serves as a contextual dimension against which structural paths are evaluated
Methods	Repeated-measures ANOVA; growth curve modeling; multilevel growth models; trend analyses; mean-level changes	Multigroup analyses (e.g., using regression models, path models, or SEM); invariance testing; result comparisons across waves
Significance	Evidence of systematic temporal change; supports developmental or adaptation-based explanations	Evidence that estimated relationships differ across time; suggests contextual contingency or temporal moderation
Nonsignificance	Suggests aggregate stability (assuming adequate power); consistent with stability or persistence	Supports invariance; indicates theoretical generalizability across time periods
Level of inference	Mean-level or trajectory-level change in variables	Stability or variability of theoretical mechanisms
Theoretical contribution	Identifies whether phenomena evolve over time	Tests robustness and temporal generalizability of theoretical models



to model change over time (Bollen and Curran 2006), this approach evaluates whether a theoretical model exhibits temporal robustness. Methodologically, the procedure is equivalent to a conventional MGA in which estimated coefficients are compared across groups. Conceptually, however, the focus differs. Whereas MGA is most often used to identify and explain differences between groups (Sarstedt et al. 2011), longitudinal MGA uses time as the grouping variable and evaluates whether theoretically specified relationships remain invariant across measurement occasions. Consequently, the absence of significant differences is not merely a lack of evidence for change but, when stability is theoretically anticipated, evidence consistent with the persistence of the underlying model across temporal contexts.

In this sense, longitudinal MGA represents a test of theoretical persistence rather than merely a technique for documenting change (Little 2013; Vandenberg and Lance 2000). When measurement invariance has been established and stability is theoretically expected, nonsignificant differences in estimated coefficients provide evidence consistent with temporal robustness rather than simply reflecting a failure to detect change. Such conclusions require several methodological conditions, including a priori hypothesis (e.g., Kerr 1998), established measurement invariance (e.g., Vandenberg and Lance 2000), and adequate sample size and statistical power (e.g., Cohen 1994; Gelman and Carlin 2014).

Hypotheses derived from the theoretical model should be specified prior to data analysis to ensure that both significant and nonsignificant findings can be meaningfully interpreted (Kerr 1998). Null effects are only informative when they correspond to theoretically grounded expectations and when null hypotheses are explicitly articulated (Cohen 1994). Moreover, interpreting nonsignificant differences as evidence of stability presupposes adequate statistical power, as underpowered designs increase the risk of failing to detect substantively meaningful differences (Cohen 1994; Gelman and Carlin 2014; Greenland et al. 2016; Hoenig and Heisey 2001). The results must therefore be evaluated considering confidence intervals and sample size.

In longitudinal analyses, these requirements are complemented by the need to establish measurement invariance. Measurement invariance ensures that latent constructs are measured equivalently across time points; without it, observed differences may reflect changes in measurement properties rather than substantive changes in constructs themselves (Vandenberg and Lance 2000; Hair et al. 2024; Henseler et al. 2016). Measurement invariance therefore constitutes a prerequisite for interpreting longitudinal MGAs. When measurement invariance holds, MGAs allow researchers to test whether models and their estimated relationships significantly differ across time points.

Once invariance is established, multigroup SEM enables the assessment of whether theoretically established relationships differ across time. While empirical research often prioritizes statistically significant results (Franco et al. 2014; McShane et al. 2019), null results in longitudinal research can provide evidence of temporal stability and support claims of generalizability when models are theory-driven (e.g., Ployhart and Vandenberg 2010; Vandenberg and Lance 2000). Interpreting null results requires an inferential framework that extends beyond conventional NHST (Wasserstein et al. 2019). A nonsignificant result alone does not imply the absence of an effect, as confidence intervals may still include effect sizes of theoretical or practical relevance (Altman and Bland 1995; Hawkins and Samuels 2021). Null results only become substantively meaningful when confidence intervals are sufficiently narrow to exclude effect sizes of substantive relevance, providing evidence for their absence in a theoretical context (Cohen 1994; Gelman and Carlin 2014). This logic is formalized in equivalence testing, which allows researchers to reject effect sizes smaller than a theoretically justified smallest effect size of interest and to draw affirmative conclusions about negligible effects (Lakens 2017).

Furthermore, the interpretability of null results depend on study design, statistical power, and the precision of estimation. More broadly, uncertainty in empirical research arises not only from sampling variability but also from measurement imprecision, model specification choices, and contextual influences that affect estimated effects (Sarstedt 2026). In underpowered studies, wide confidence intervals limit interpretability, as they may still encompass meaningful effects. Increasing sample size may reduce this uncertainty and narrow confidence intervals, enabling more decisive conclusions, provided that estimated effects remain small. In this sense, null results are not defined by nonsignificance per se, but by the degree to which the data credibly rule out effect sizes predicted by competing theoretical explanations (Cohen 1994; Greenland et al. 2016; Lakens 2017; Landis et al. 2014; McShane et al. 2019).

Empirical illustration

Model

To demonstrate the significance of null results, we use an empirical illustration that examines the influence of motivational, social, and work-related characteristics on employees' turnover intentions. Turnover intention is a key outcome variable in applied contexts such as employer branding, personnel marketing, and employee experience analytics, where understanding drivers of employee retention is



of central importance. We draw on the adapted turnover intention model Ehrlich (2024) proposed, which estimates the effects of motivational work characteristics (MWC), met expectations (ME), and leader-membership exchange (LMX) on turnover intentions (TI), with affective commitment (AC) and job satisfaction (JS) specified as mediating mechanisms (Fig. 1). We used established measures for our key constructs (Ehrlich 2024; Table 4 in the appendix provides an overview).

Turnover intention has long been conceptualized as a central cognitive antecedent of employee turnover and a key outcome variable in organizational and management research (Hom et al. 2017; Mobley 1977). Classic models emphasize job satisfaction and organizational commitment as primary drivers of employees' withdrawal cognitions (Meyer and Allen 1991), while subsequent research has shown that work characteristics, leadership, and the alignment between expectations and experienced job conditions play additional, theoretically distinct roles (Hackman and Oldham 1976; Podsakoff et al. 2007). Together, this literature implies that turnover intention reflects a relatively stable evaluative state that emerges from multiple interrelated attitudinal and contextual inputs, rather than a purely reactive or short-lived response to isolated events. Building on this research, Ehrlich (2024) developed a theoretically integrative model of turnover intention that embeds core attitudinal constructs, such as job satisfaction and organizational commitment, in a broader framework incorporating expectations, task-related perceptions, and leadership influences (for the hypothesized relationships in the model, see Ehrlich 2024). Rather than treating these factors as competing explanations, the model conceptualizes them as jointly shaping employees' intention to leave through stable cognitive evaluations of their work situation.

While the theoretical model builds on this established framework, this research's primary contribution does not lie in extending the model, but in its methodological perspective—specifically, in interpreting longitudinal comparisons and the absence of change as informative null results. This structure reflects a large body of research suggesting that turnover intentions are not inherently volatile but often reflect equilibrium states shaped by enduring features of the job and the organization (Hom et al. 2017; Steel and Lounsbury 2009). Ehrlich's (2024) model lends itself naturally to longitudinal analysis in which stability itself is theoretically meaningful. Many turnover theories do not imply unconditional persistence, but rather stability under relatively stable contextual conditions: once formed, turnover intentions are expected to remain largely stable unless fundamental aspects of the work environment shift (Hom et al. 2017; Mobley 1977). Accordingly, observing no systematic change in the model's theoretically established relationships over time should not be interpreted as direct confirmation of the underlying theory, but rather as being consistent with theoretical expectations of temporal stability under the conditions examined. We therefore use Ehrlich's (2024) model as an empirical illustration of how longitudinal data can generate substantively relevant insights even when estimated effects remain stable across time, demonstrating that the absence of change can still provide meaningful information for theory and practice.

Data and method

The empirical estimation of the model presented in Fig. 1 is based on data from the Swiss Human Resource Barometer (SHRB) collected biannually in five waves (2012, 2014, 2016, 2018, and 2020; Grote et al. 2013, 2017; Pfrombeck et al. 2018, 2020; Staffelbach et al. 2022). The SHRB is a

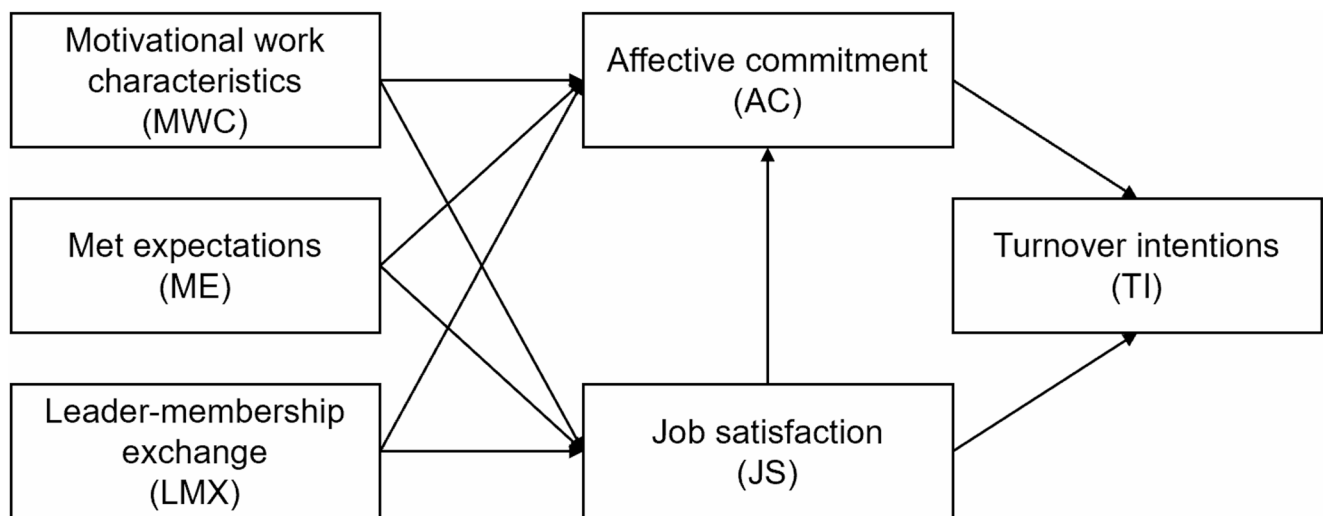


Fig. 1 Research model



collaborative research initiative of the University of Zurich, ETH Zurich, and the University of Lucerne. The sample data for each wave were provided by the Swiss Federal Statistical Office (FORS) and comprises 1,483 observations in 2012, 1,386 in 2014, 1,504 in 2016, 1,947 in 2018, and 1,919 in 2020. FORS ensured that each sample is representative of Swiss workforce at the time of data collection. The datasets are available through the SHRB data providers, subject to their access and usage conditions. Specifically, the 2012, 2014, 2016, 2018, and 2020 datasets used in this research are distributed through the SWISSUbase (see also Ehrlich 2024).¹

This high level of sample representativeness constitutes a critical quality criterion of the SHRB data and supports the generalizability of the empirical analyses and findings derived from five waves. More specifically, the period from 2012 to 2020 was marked by several political, economic, and societal developments in Switzerland that may have affected employees' work-related evaluations and turnover intentions.² Rather than presuming that such contextual developments changed the estimated relationships, we explicitly test whether the theoretically established relationships remain stable across waves. Evidence of stability would suggest that the model captures mechanisms that persist across contextual fluctuations. Evidence of change, in contrast, would identify temporal boundary conditions and refine theoretical understanding.

Note. Most researchers would classify this design as a repeated cross-sectional study. Temporality is however central to the research question and constitutes the primary object of analysis, which opens room for considering whether such designs may, under certain conditions, be conceptualized as longitudinal in the social sciences. Drawing on the definition by Ployhart and Vandenberg (2010), longitudinal research requires repeated measurements over time from the same units of observation. Although the individual respondents differ across waves in this study, each sample is representative of the Swiss population, and the analytical focus is not on individual-level trajectories but on changes at the population level. From this perspective, the Swiss

population can be conceptualized as the unit of analysis, such that repeated measurements over time are obtained for the same underlying entity. Accordingly, we argue that the design exhibits key characteristics of longitudinal research, as it captures temporally ordered change in a consistent and linkable manner at the population level.

Following Ehrlich (2024), the five datasets were prepared by harmonizing the variables across waves so that only consistently formulated and consistently coded items were used. Variables not required for the analysis were removed, the negatively worded turnover intention item was recoded, and the met expectations indicators were computed as difference scores between employer expectations and perceived employer offerings. Following Hair et al. (2027), we removed 166 out of 8,330 observations across the five samples because they contained more than 15% missing values. The deletion of approximately 2% of the overall data does not considerably affect the sample's representativeness. For the remaining missing values, we used mean replacement, which offers robust outcomes in our analysis (e.g., Hair et al. 2027; Grimm and Wagner 2020). Since the number of remaining missing data points is relatively small, with 4,584 out of 158,270 values missing, corresponding to 2.9%, alternative missing data treatment options (e.g., Wang et al. 2022) do not show noticeable differences in the results. A detailed description of the data preparation and operationalization is provided in Ehrlich (2024).

We assessed the potential influence of common method variance before estimating the research model, given that the data are based on survey responses. Following the marker-variable logic discussed by Tehseen et al. (2017) and the item-level correction approach described by Merkle (2025), we used the item "Spend what I have and enjoy it—Save for the future" (i.e., measured on a nine-point scale) as the theoretically most unrelated marker variable available across all five survey years. This item captures a general financial orientation rather than work-related attitudes or evaluations. The correlations between this marker variable and the items used in the research model were consistently below 0.1 across all five survey years, supporting its suitability. The item-level method variance extracted was 0.2% or lower in each year, indicating that common method variance is unlikely to materially affect the findings.

In this research, we used structural equation modeling (SEM) for the model estimation and results analysis. More specifically, since we were particularly interested in estimating a model that includes both reflective and formative constructs (e.g., Coltman et al. 2008; Diamantopoulos and Winklhofer 2001) and analyzing the model's predictive capabilities (e.g., Legate et al. 2023, 2024), we applied partial least squares SEM (PLS-SEM; Hair et al. 2027; Lohmöller 1989; Wold 1982) using the SmartPLS 4 software for

¹ The datasets used in this research are available through FORS/SWISSUbase subject to the applicable access and usage conditions. In accordance with the data use agreement, the data are used exclusively for scientific research, are reported only in aggregated form, and no information is disclosed that would allow the identification of individual respondents. The datasets are cited according to the data providers' recommended citation format (see references).

² For example, the 2014 initiative against mass immigration introduced considerable policy uncertainty regarding labor market access and future employment conditions, while the 2015 appreciation of the Swiss franc created economic pressure for internationally oriented firms. In 2020, the COVID-19 pandemic further disrupted the Swiss labor market, including through sharp increases in short-time work arrangements.



data analysis (Ringle et al. 2024). To assess the model's out-of-sample predictive capabilities, we applied the cross-validated predictive ability test (CVPAT; Liengaard et al. 2021; Sharma et al. 2023). To assess potential differences across survey years, we conducted a multigroup analysis (MGA; Chin and Dibbern 2010) within the PLS-SEM framework, following the methodological guidelines Hair et al. (2024) proposed.

As measurement invariance is a prerequisite for meaningful group comparisons, we first evaluated invariance using the MICOM procedure in PLS-SEM (Henseler et al. 2016; Liengaard et al. 2021). After establishing measurement invariance, we performed the MGA using the permutation approach (Chin and Dibbern 2010; Hair et al. 2024). This approach is particularly appropriate given our objective of testing whether the estimated model parameters remain statistically equivalent across survey waves. By doing so, the analysis enables an assessment of the robustness and temporal stability of the theoretically derived relationships among the key antecedents of turnover intention.

Results

The presentation and evaluation of the PLS-SEM results follow established recommendations and methodological guidelines (e.g., Guenther et al. 2023; Hair et al. 2024, 2027; Ringle et al. 2020, 2023; Sarstedt et al. 2025). The assessment began with evaluating the measurement models. Almost all indicator loadings of the reflective constructs (i.e., affective commitment, leader-membership exchange, and turnover intention) exceeded the recommended threshold of 0.7, supporting indicator reliability (Table 5 in the appendix).³ In addition, Cronbach's α , the composite reliability ρ_A , and the composite reliability ρ_C values were

above 0.708, indicating satisfactory internal consistency reliability for the reflective constructs (Table 6 in the appendix). Convergent validity was supported, as all average variance extracted (AVE) values exceeded the critical threshold of 0.5 (Table 6 in the appendix). Discriminant validity was further established using the heterotrait-monotrait ratio (HTMT; Henseler et al. 2015), with all values falling significantly below the conservative cutoff of 0.85 (Table 7 in the appendix; see also Franke and Sarstedt 2019; Hair et al. 2027).

For the formatively measured constructs, collinearity was assessed using the variance inflation factor (VIF). All VIF values were below three, indicating that collinearity was not a concern (Table 8 in the appendix). Most indicator weights of the formative constructs were statistically significant ($p < 0.05$; Table 9 in the appendix). Exceptions were PV2_diff, which was nonsignificant across almost all time points except 2018, PV4_diff, which was nonsignificant across almost all time points except 2016, and PV5_diff, which was nonsignificant in 2014. However, nonsignificant formative indicator weights should not automatically be removed, because formative indicators capture distinct facets of the construct and their removal may compromise content validity. Moreover, when formative indicators have nonsignificant weights but statistically significant outer loadings, they can still contribute to the absolute meaning of the construct and may therefore be included (Hair et al. 2019, 2027). Arguably the indicator could have been removed, yet we retained these indicators due to their statistically significant outer loadings, their contribution to the constructs' content validity, and their role in supporting measurement consistency over time.

The structural model's results evaluation started with the collinearity assessment. The highest VIF in any of the structural models from 2012 to 2020 had a value of 1.757, which was clearly below the cutoff point of 3. We therefore concluded that collinearity was not critical in our empirical analyses. Model fit was assessed with the standardized root mean square residual (SRMR), which was below 0.08 for all model estimations, therefore model fit was deemed acceptable across all measurement time points (e.g., Dash and Paul 2021; Hair et al. 2027; see also Henseler et al. 2014). Whereas Table 2 presents the R^2 values for the endogenous constructs, which are at satisfactory levels (i.e., the lowest R^2 has a value of 0.293; see also Ehrlich 2024), Fig. 2 reports the path coefficients for all five measurement years, all of which are statistically significant (see also Table 10). The signs and magnitudes of the coefficients are largely consistent over time, providing initial evidence that the theoretically established relationships in the model are stable across waves. Moreover, all indirect effects are statistically significant across the five survey years (Table 11), indicating

Table 2 R^2 values

Constructs	2012	2014	2016	2018	2020
Job satisfaction (JS)	0.431	0.389	0.396	0.361	0.313
Affective commitment (AC)	0.361	0.375	0.327	0.381	0.377
Turnover intention (TI)	0.314	0.315	0.294	0.334	0.293

³ While the third indicator of turnover intention (ti3) showed satisfactory loadings in the survey years 2016 to 2020, it exhibited loadings around 0.6 in the survey years 2012 and 2014. However, indicators with loadings between 0.40 and 0.708 should not be removed automatically. Rather, indicator removal should only be considered when it substantially improves internal consistency reliability or convergent validity, that is, when it helps the construct meet the relevant evaluation criteria, and when deletion does not compromise content validity (e.g., Hair et al. 2027). In our case, removing ti3 in 2012 and 2014 was not warranted because the other items of the construct showed strong loadings above 0.8 and the reliability and validity criteria for the construct exceeded the established cutoff values across all years. Retaining the indicator therefore supported measurement consistency over time.





Fig. 2 Model with path coefficients

that affective commitment and job satisfaction consistently function as mediating mechanisms in the model. The results therefore show that the antecedents influence turnover intention not only directly but also indirectly through the attitudinal constructs included in the model. Finally, Table 12 reports the total effects, which are particularly relevant from a practical perspective because they capture the overall influence of the antecedent constructs on turnover intention through all direct and indirect pathways. The total effects of leader-membership exchange, met expectations, and motivational work characteristics on turnover intention are consistently negative and statistically significant across all five survey years. Overall, the direct, indirect, and total effect results jointly support the practical relevance and temporal stability of the model.

Given the purpose of this empirical example, which is to illustrate the interpretation and relevance of null effects in longitudinal data, we do not further discuss the results of the individual cross-sectional models. Instead, we focus next on the CVPAT analysis, which assesses the model's predictive capabilities, and then on the multigroup analysis, which directly assesses whether the estimated relationships differ across measurement occasions.

To support the model's predictive capabilities, the CVPAT's average loss value of the PLS-SEM results should be lower than the average loss value obtained from the more naïve IA prediction benchmark and the more conservative LM prediction benchmark. In comparison to the IA (indicator average) benchmark, we found that PLS-SEM had better

predictive capabilities across all years, as indicated by the negative and significant average loss differences (Table 13 in the appendix). For the LM (linear model) benchmark, the average loss differences in the overall model were all positive, but very small and nonsignificant (Table 14 in the appendix). PLS-SEM's predictive capabilities were therefore not significantly better than those of the more conservative LM benchmark. In summary, the CVPAT outcomes indicate that the model has low to moderate predictive capabilities, as the PLS-SEM results consistently and statistically significantly outperformed the IA benchmark but did not surpass the LM benchmark (Liengard et al. 2021; Sharma et al. 2023).

This empirical illustration aims to examine whether there is a robustness of the model over time. For this purpose, we carried out a PLS-SEM multigroup analysis (MGA; (Chin and Dibbern 2010; Hair et al. 2024). Instead of applying the MGA to different groups in a cross-sectional dataset, we used the MGA for a longitudinal analysis in which every point in time represents a group of its own (e.g., Roemer 2016). In longitudinal multigroup analyses, group comparisons should be guided by theoretical expectations rather than conducted exhaustively (e.g., Henseler et al. 2016).

Prior research emphasized that once measurement invariance is established, only substantively meaningful comparisons are required to address the research question (Vandenberg and Lance 2000; Vandenberg 2002). Following longitudinal modeling principles, we therefore focused on adjacent-year comparisons to assess incremental change

(t to $t+1$) and a first-to-last comparison to evaluate long-term stability (Singer and Willett 2003). Testing all possible pairwise combinations would not provide additional theoretical insight but would substantially increase the number of statistical tests without corresponding interpretive value.

We assessed measurement invariance using the MICOM procedure (Henseler et al. 2016). Partial measurement invariance is sufficient because our analysis focuses on standardized coefficients and the datasets are not pooled (Hair et al. 2024). Step 1 of the MICOM procedure, configural invariance, was established because the same indicators, data treatment, model specification, and algorithm settings were used for the model estimations across all years. Step 2, compositional invariance, was also established, as all permutation-based p values were above 0.05. The MICOM results therefore indicate partial measurement invariance for all comparisons, allowing us to proceed with the MGA analyses (see Ehrlich 2024, for a more detailed discussion of the results).

Next, we focus on the results of the permutation-based MGA (Chin and Dibbern 2010). Table 15 (in the appendix) shows the differences between the path coefficients for 1,000 permutations and a 95% confidence interval in a two-sided test. In pairwise comparisons across the different years, we examined 45 path coefficients for significant differences (for further details, see Ehrlich 2024). To account for α -inflation resulting from multiple testing, we applied the Bonferroni correction, as recommended by Hair et al. (2024); alternatively, researchers may use the Šidák correction. Using the standard global significance level of $\alpha=0.05$ across the 45 path coefficient comparisons, the local significance level was adjusted to approximately $\alpha=0.001$. Based on this correction, only one path coefficient difference was clearly significant: the LMX $\rightarrow$ AC relationship differed between 2012 and 2014. Three additional comparisons yielded permutation p values of 0.001 and therefore represent borderline cases: the LMX $\rightarrow$ AC relationship between 2014 and 2016, the AC $\rightarrow$ TI relationship between 2012 and 2020, and the JS $\rightarrow$ TI relationship between 2012 and 2020. These isolated significant and borderline differences may be interpreted considering broader contextual developments during the observation period.⁴ “Apart from the one clearly significant

comparison and the three borderline cases, the remaining 41 of the 45 direct effect comparisons were nonsignificant. Relatedly, all total indirect effects and all total effects in the model showed nonsignificant differences across time, further supporting the temporal stability of the model (for more detailed results and discussion, see Ehrlich 2024).

Discussion

The aim of this empirical illustration was to display a scenario with real longitudinal data where the primary outcomes are null results but with high theoretical and empirical relevance. Beyond the substantive implications of the empirical illustration for research on turnover intention (for more details, see Ehrlich 2024), the findings carry an important methodological insight: the presence of substantively relevant null results in longitudinal data analysis. Across all multigroup comparisons spanning five measurement occasions, only one (and three borderline) statistically significant differences out of 45 tested structural relationships were observed. Rather than indicating an absence of information, these results provide evidence that the underlying model structure is stable over time. In longitudinal research, such results directly speak to theoretical claims about persistence, stability, and institutionalization, and therefore constitute a meaningful empirical outcome rather than an inconclusive one (Hom et al. 2017; Vandenberg and Lance 2000). The results of the empirical illustration based on Ehrlich’s (2024) adapted turnover intention model illustrate how longitudinal data can yield substantively relevant insights even when estimated effects remain stable over time. Across all five measurement occasions, the structural paths in the model are statistically significant, indicating that the relationships among motivational work characteristics (MWC), met expectations (ME), and leader-membership exchange (LMX) on turnover intentions (TI), with affective commitment (AC) and job satisfaction (JS) specified as mediating mechanisms, are consistently supported.

Considering the CVPAT analysis, the results indicate that the model has low to moderate predictive power across all time points and that the predictive power does not change across time points. These patterns suggest that the model’s underlying theoretical structure is robust and that the predictive power is stable across each observed time point, which builds the foundation for being able to interpret null results of the MGA analysis to assess temporal change. For the comparison of path coefficients across time, only one difference was clearly statistically significant after applying the Bonferroni correction whereas three additional comparisons yielded borderline p values. Importantly, even when these borderline cases are considered, the significant or

⁴ As noted, the 2012–2020 period encompassed several political, economic, and societal events in Switzerland that may have influenced employees’ evaluations of their work environment and turnover-related attitudes. For example, the differences involving the LMX $\rightarrow$ AC relationship may indicate temporary contextual sensitivity in the role of leadership relationships for employees’ affective attachment to the organization. Likewise, the differences involving the AC $\rightarrow$ TI and JS $\rightarrow$ TI relationships may be consistent with changing labor market conditions toward the end of the observation period, like the beginning of the COVID-19 pandemic. These interpretations remain speculative, however, as the research design does not permit causal attribution of coefficient differences to specific contextual events.



near-significant differences remain limited to a small number of selected paths and do not indicate broad temporal instability of the model. The overwhelming majority of path coefficient comparisons were nonsignificant, supporting the conclusion that the model's structural relationships are largely stable across the observed time points.

Prior methodological work cautioned against interpreting sporadic significant differences as evidence of substantive change, particularly in longitudinal settings where stable processes are theoretically expected (Meehl 1978; Singer and Willett 2003; Vandenberg and Lance 2000). When the differences in path coefficients remain small, confidence intervals largely overlap, and no coherent temporal trajectory emerges, isolated significant MGA results are more plausibly attributable to random fluctuation than to meaningful changes in underlying mechanisms (Gelman and Carlin 2014; McShane et al. 2019). Taken together, the predominance of nonsignificant differences and the absence of systematic instability provide affirmative evidence for temporal robustness of the model, reinforcing the substantive relevance of stability rather than change. While null results are often omitted or dismissed as uninformative (e.g., Franco et al. 2014), in this context they are of substantial theoretical and practical relevance. The absence of significant differences indicates that the magnitude of the modeled relationships remains stable over an eight-year period, providing affirmative evidence of temporal robustness (e.g., Little 2013; Vandenberg and Lance 2000). Rather than signaling a lack of insight, the nonsignificant MGA results meaningfully support theories by adding generalizability.

Guidelines

The results of this empirical illustration suggest several guidelines for the reporting and interpretation of null results in longitudinal research. Researchers should specify *ex ante* whether stability or change is expected and treat null differences across measurement occasions as explicit tests of these expectations. When theory predicts persistence, equilibrium, or stable relationships, finding no statistically significant differences in the estimated path coefficients across waves should be interpreted as evidence consistent with temporal stability rather than as a failure to detect an effect (Vandenberg and Grelle 2010). The inferential task in longitudinal research is therefore not limited to identifying change; it also includes assessing whether theoretically meaningful change is absent.

As Judd and Kenny (1981) note, stability over time can disconfirm dynamic mechanisms just as decisively as

observed change can support them. When secondary data are used, researchers have limited control over the timing, spacing, and duration of measurement waves. Consequently, short-term dynamics or temporary fluctuations may remain unobserved, particularly when the original data collection was not designed around the focal theoretical question, as in our empirical illustration (e.g., Bell et al. 2022; Johnston 2014). At the same time, large-scale secondary datasets compiled by professional data institutes often provide broad temporal coverage and substantial sample sizes, which may still allow researchers to detect more durable and systematic patterns (e.g., Bell et al. 2022; Johnston 2014). Nevertheless, extended field periods can also blur short-term processes. Researchers should therefore state clearly which forms of temporal dynamics their data can capture and which may remain unobserved. For studies based on primary data collection, the spacing of measurement occasions should be determined on theoretical grounds so that the temporal design can detect the processes under investigation (Ployhart and Vandenberg 2010). Longitudinal analyses should model time explicitly and avoid selectively emphasizing only those periods that yield statistically significant results.

When using PLS-SEM to analyze longitudinal data, it is important to assess the measurement models and structural models separately for each time point. Only after the measurement models fulfill established reliability and validity criteria and the structural models meet the relevant evaluation criteria at each wave, can meaningful comparisons across time be conducted. The same logic applies to the evaluation of predictive capabilities. A CVPAT analysis should be conducted for each time wave to assess the model's out-of-sample predictive performance. If the CVPAT results indicate comparable predictive performance across waves, it provides additional evidence that the model's predictive ability is stable over time. The predictive orientation of PLS-SEM together with the CVPAT analysis can strengthen the interpretation of longitudinal stability. Because PLS-SEM emphasizes explained variance and predictive relevance, researchers can examine whether stability in structural coefficients is accompanied by stability in predictive performance across waves (Shmueli et al. 2019; Sharma et al. 2023). When both the path estimates and the predictive capabilities remain comparable over time, the argument for structural stability gains additional substantive support.

Any claim regarding stability or change in estimated relationships additionally depends on evidence of at least partial measurement invariance across time. Without evidence that constructs are measured equivalently at each wave, apparent stability in structural coefficients cannot be



attributed confidently to substantive processes rather than to measurement artifacts (Vandenberg and Lance 2000; Vandenberg 2002). Researchers should therefore routinely assess and report measurement invariance and explicitly connect the results to their interpretation of stable or null structural differences. Importantly, partial measurement invariance is sufficient for MGA when the objective is to compare standardized structural coefficients across groups or time points and the datasets are not pooled (Henseler et al. 2016; Hair et al. 2024). In the MICOM procedure, partial measurement invariance requires configural invariance and compositional invariance. Configural invariance means that the same indicators, data treatment, model specification, and algorithm settings are used across groups, whereas compositional invariance indicates that the composites are formed similarly across groups. Full measurement invariance additionally requires equality of composite means and variances. This stricter level is mainly required when researchers intend to pool data or compare group-specific means and variances, but it is not required for the MGA conducted in this research. In our empirical illustration, we assessed measurement invariance using the MICOM procedure, which is widely used in PLS-SEM (e.g., Henseler et al. 2016). In covariance-based SEM, multigroup confirmatory factor analysis (MGCFA; e.g., French and Finch 2008) and the approach by Diamantopoulos and Papadopoulos (2010) represent suitable alternatives.

Once at least partial measurement invariance has been established, MGA enables researchers to estimate the structural model separately for different time points and to formally test whether path coefficients differ across waves. When longitudinal hypotheses imply consistency, nonsignificant between-wave differences in path coefficients can be interpreted as evidence consistent with structural stability. For this purpose, we employed the permutation-based MGA, which represents an adequate approach for testing the significance of coefficient differences across time (Chin and Dibbern 2010). Alternative approaches include the t-test, under the assumption of equal or unequal variances, and the bootstrap-based MGA (for an overview and further explanations, see Hair et al. 2024). To avoid reliability and statistical power issues in the comparison, permutation-based MGA should not be applied to extremely unbalanced sample sizes (Hair et al. 2024). In our empirical illustration, the sample sizes across waves are reasonably balanced (i.e., with 1,483 observations in 2012, 1,386 in 2014, 1,504 in 2016, 1,947 in 2018, and 1,919 in 2020). Researchers may also consider using different approaches to substantiate their findings across methods in a multimethod MGA approach (Sarstedt et al. 2017).

Finally, null results should be reported with the same degree of transparency as statistically significant results. Researchers should present point estimates, standard errors, confidence intervals, and other indicators of estimate precision rather than relying exclusively on dichotomous significance decisions (Gelman and Carlin 2014). Where possible, they should also specify a smallest effect size of interest so that readers can assess whether the observed absence of differences is merely statistically nonsignificant or also substantively negligible. Such reporting enables readers to evaluate which magnitudes of change can reasonably be ruled out and whether the evidence is consistent with meaningful temporal stability. Table 3 provides these established guidelines as a checklist for researchers to follow when conducting research with expected null results.

Conclusion

Historically, several authors have cautioned against interpreting null results, particularly when relying on conventional null hypothesis significance testing (NHST; Cohen 1994; Gelman and Carlin 2014; McShane et al. 2019; Meehl 2004). From this perspective, failure to detect differences is often framed as inconclusive due to potential issues of power, measurement error, or model misspecification (e.g., Cohen 1994; Nickerson 2000). As a result, nonsignificant findings have frequently been treated as non-results and discounted in empirical research. More recent methodological literature explicitly challenges this view (Kozlov 2024; Landis et al. 2014), emphasizing that null results can be interpretable and informative when estimates are sufficiently precise and theoretically grounded (e.g., Gelman and Carlin 2014; Lakens et al. 2020). This perspective is particularly relevant in longitudinal research, where stability rather than change may be the theoretically expected outcome. When theories predict persistent relationships, the absence of statistically significant differences across time does not weaken inference but instead provides evidence that the proposed mechanisms operate consistently under varying conditions (e.g., Shadish et al. 2002; Vandenberg and Grelle 2010; Vandenberg and Lance 2000). A theoretically established model that exhibits substantial fluctuations in estimated relationships across time points may therefore be of limited value for researchers and practitioners. Conversely, when supported by appropriate research design and evidence of measurement invariance, longitudinal null effects can provide informative evidence for the stability and robustness of theoretical relationships.



Table 3 Guidelines for supporting and reporting longitudinal PLS-SEM results

	Recommendations	Key references
Theoretical expectations	- Specify ex ante whether stability or change is expected - Treat null differences as explicit tests of theoretical expectations - Interpret nonsignificant differences as evidence of stability when theory predicts persistence	Judd and Kenny (1981), Vandenberg and Grelle (2010)
Research design	- Clearly define which temporal dynamics can be captured - For primary data, space measurement occasions based on theoretical considerations - Acknowledge limitations of secondary data (timing, wave spacing, short-term dynamics)	Ployhart and Vandenberg (2010), Bell et al. (2022), Johnston (2014)
Model assessment	- Evaluate measurement and structural models separately at each wave - Ensure that reliability and validity criteria are met before comparisons across waves - Report relevant measurement model assessment criteria results	Hair et al. (2027), Sarstedt et al. (2025)
Predictive evaluation	- Conduct CVPAT or equivalent predictive assessments for each wave - Assess whether predictive performance remains comparable across waves - Use stability in predictive performance as additional evidence supporting structural stability	Shmueli et al. (2019), Sharma et al. (2023)
Measurement invariance	- Assess and report at least partial invariance across waves before interpreting structural differences - In PLS-SEM, establish at least partial measurement invariance before conducting MGA based on standardized coefficients - Use MICOM in PLS-SEM or MGCFA in CB-SEM to assess measurement invariance	Vandenberg and Lance (2000), Henseler et al. (2016), French and Finch (2008), Diamantopoulos and Papadopoulos (2010)
Multigroup analysis (MGA)	- Test whether path coefficients differ across waves using permutation-based, bootstrap-based, or t-test MGA methods - Ensure that group sizes are not extremely unbalanced, particularly when applying permutation-based MGA - Correct for multiple testing when several comparisons are conducted across waves - Interpret nonsignificant results as evidence consistent with structural stability when stability is theoretically expected	Chin and Dibbern (2010), Hair et al. (2024), Sarstedt et al. (2017)
Direct, indirect, and total effects	- Report direct, indirect, and, where relevant, total effects to capture both theoretical mechanisms and practical implications - Interpret total effects when the overall influence of antecedents on key outcomes is relevant - Assess whether direct, indirect, and total effects remain stable across waves	Hair et al. (2027), Memon et al. (2026)
Transparent reporting	- Report estimated coefficients, confidence intervals, and p values for relevant structural relationships and comparisons across waves - Report nonsignificant findings with the same level of transparency as significant findings - Clearly distinguish between statistically nonsignificant differences and evidence consistent with theoretical expectations of stability - Document data quality, sample characteristics, missing data handling, and, where applicable, the use of weighting variables - In general, follow state-of-the-art open science practices	Adler et al. (2023b), Gelman and Carlin (2014)



In consequence, on a broader scope, our research also contributes to moving toward open science in marketing research (Deer et al. 2025). Central goals of the open science movement are to enhance transparency, reproducibility, and the cumulative evaluation of empirical claims by making research designs, data, and analytical decisions accessible and testable (e.g., Nosek et al. 2015). Reform efforts have highlighted how publication bias and selective reporting have distorted the evidentiary value of statistically significant findings (Ioannidis 2005; Simmons et al. 2011). Researchers therefore continuously enhance the toolbox to evaluate the trustworthiness of published findings (e.g., Adler et al. 2023a).

Our longitudinal design—assessing the same theoretical model across five independent time points—can be understood as a form of structured temporal replication. We therefore subscribe to frequent calls that demand replication of published research, which also represents a particularly important concern in SEM (e.g., Adler et al. 2023a, 2023b), as well as to general calls for transparency, reproducibility, and open access to data in research publications (American Psychological Association 2020; Nosek et al. 2015). Greater transparency also facilitates the evaluation of data quality and encourages researchers to document sampling procedures, measurement decisions, and data-processing steps more rigorously. This becomes increasingly important in an era characterized by large-scale digital datasets and AI-supported analytics, where the validity of scientific conclusions depends fundamentally on the quality and traceability of the underlying data. Open science practices therefore contribute not only to reproducibility but also to a more transparent assessment of the strengths and limitations of empirical evidence. By explicitly testing the model's stability over time and transparently reporting nonsignificant multigroup differences, the study aligns with the core principles of the open science community: allowing findings to be re-evaluated, reproduced, and scrutinized across contexts. When hypotheses are specified *a priori* and statistical power is considered, null results across time do not signal a lack of contribution but instead provide evidence for the robustness and generalizability of the proposed theoretical mechanisms. As large-scale replication initiatives have demonstrated (Open Science Collaboration, 2015), cumulative science advances not only through detecting change, but also through demonstrating stability under varying historical conditions.

Future research may aim at replicating this study's results using the outcomes of additional model estimation methods such as CB-SEM, GSCA, and sum scores regression

in a multimethod approach to SEM (e.g., Cheah et al. 2026; Guenther et al. 2025; Hair et al. 2026; Margalina et al. 2026; Sarstedt et al. 2020a, 2024). Thereby, researchers further support the robustness of the presented findings and conclusions based on the estimated model. In addition, researchers may consider robustness checks (Sarstedt et al. 2020b; Vaithilingam et al. 2024), including, for instance, nonlinear relationships (e.g., Basco et al. 2021; Hair et al. 2024), unobserved heterogeneity (e.g., Becker et al. 2013; Sarstedt et al. 2022b), multigroup analyses (e.g., Chin and Dibbern 2010; Hair et al. 2024), and the consistency of these findings across time. Furthermore, researchers may also consider the application of extended techniques such as importance performance map analysis (IPMA; e.g., Hair et al. 2024; Ringle and Sarstedt 2016), necessary condition analysis (NCA; e.g., Dul 2016; Becker et al. 2026), and treating endogeneity issues (Becker et al. 2022; Liengaard et al. 2025). Moreover, future research may further refine the model by examining more focused, theoretically grounded frameworks or broader conceptual approaches, provided the underlying data supports such extensions. Additionally, the SHRB dataset includes special-topic items in each wave that could be incorporated into the respective year, where theoretically justified. Integrating these items may help illuminate the relevance of specific topics within the overall models. Given the stability of effects observed in this research, it is reasonable to expect similar patterns to emerge in other years, even in the absence of special-topic data. Predictive model comparisons—such as those proposed by Liengard et al. (2021)—may offer valuable guidance for future model extensions by evaluating the performance of the proposed model against alternative models. Finally, the future research directions presented here for our illustrative example may be generalized into a broader framework. For example, the EP-mix framework Sharma et al. (2024) proposed to enhance research replication could be extended by incorporating the robustness of model results across time. In this context, researchers may consider extending this framework and the robustness of model estimations across time by Sarstedt's (2026) region of practical equivalence (ROPE) approach that introduces metrological uncertainty as a broader framework for interpreting null results.

Appendix

See Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15.



Table 4 Constructs and measurement models

Construct	Short	Items	Measurement	Source
Leader-membership-exchange (LMX)	lmx1	My supervisor understands my problems and needs	Reflective	Schyns (2002)
	lmx2	My supervisor would use their power to help me solve problems in my work		
	lmx3	I have an effective working relationship with my supervisor		
Met expectations (ME)		The differences between what I expect from my employer and what my employer offers me.	Formative	Raeder et al. (2009)
	PV1_diff	.. interesting job content		
	PV2_diff	.. job security		
	PV3_diff	.. loyalty		
	PV4_diff	.. opportunities to take on personal responsibility		
	PV5_diff	.. opportunities to use my skills in various ways		
	PV6_diff	.. development opportunities within the company		
	PV7_diff	.. appropriate compensation for my work		
Motivational work characteristics (MWC)	tasksign	The results of my work are likely to significantly affect the lives of other people	Formative	Morgeson and Humphrey (2006)
	taskident	The job involves completing a piece of work that has an obvious beginning and end		
	feedback	The job provides feedback on my performance		
Job satisfaction (JS)	js	How satisfied are you with your job on a scale of 1 to 10, where 1 means completely dissatisfied and 10 means completely satisfied?	Single item	Robinson and Wolfe Morrison (2000)
Affective commitment (AC)	comm1	I feel like part of the family at my organization	Reflective	Meyer et al. (1993)
	comm2	I feel emotionally attached to this organization		
	comm3	I have a strong sense of belonging to my organization		
Turnover intention (TI)	ti1	I will probably look for a new job in the near future	Reflective	Mowday et al. (1979)
	ti2	At the present time, I am actively searching for another job at a different organization		
	ti3	I am not planning to leave my current job (<i>reversed item</i>)		



Table 5 Loadings of reflective measurement models

Construct	Indicator	Loadings				
		2012	2014	2016	2018	2020
Affective commitment (AC)	com1	0.874	0.869	0.885	0.886	0.882
	com2	0.907	0.903	0.921	0.911	0.886
	com3	0.927	0.920	0.931	0.929	0.917
Leader-membership exchange (LMX)	lmx1	0.923	0.890	0.886	0.894	0.895
	lmx2	0.888	0.874	0.879	0.871	0.858
	lmx3	0.913	0.891	0.870	0.888	0.874
Turnover intention (TI)	ti1	0.917	0.915	0.917	0.918	0.892
	ti2	0.879	0.884	0.870	0.889	0.859
	ti3	0.591	0.667	0.738	0.744	0.730

Table 6 Reflective measurement model evaluation results

Construct	Criterion	2012	2014	2016	2018	2020
Affective commitment (AC)	Cronbach's α	0.886	0.879	0.899	0.895	0.876
	Composite reliability ρ_A	0.888	0.881	0.901	0.897	0.880
	Composite reliability ρ_C	0.930	0.925	0.937	0.934	0.923
	AVE	0.815	0.805	0.832	0.826	0.801
Leader-membership exchange (LMX)	Cronbach's α	0.894	0.862	0.852	0.861	0.848
	Composite reliability ρ_A	0.897	0.863	0.852	0.863	0.852
	Composite reliability ρ_C	0.934	0.915	0.910	0.915	0.908
	AVE	0.824	0.783	0.772	0.782	0.767
Turnover intention (TI)	Cronbach's α	0.726	0.769	0.798	0.812	0.770
	Composite reliability ρ_A	0.810	0.829	0.833	0.842	0.788
	Composite reliability ρ_C	0.846	0.867	0.882	0.889	0.868
	AVE	0.654	0.688	0.714	0.729	0.688

AVE average variance extracted

Table 7 Heterotrait-monotrait ratio of correlations (HTMT) results

	Job satisfaction (JS)	Leader-membership exchange (LMX)	Affective commitment (AC)
Leader-membership exchange (LMX)	2012: 0.595 (P ₉₅ : 0.629) 2014: 0.568 (P ₉₅ : 0.607) 2016: 0.585 (P ₉₅ : 0.621) 2018: 0.533 (P ₉₅ : 0.569) 2020: 0.515 (P ₉₅ : 0.551)		
Affective commitment (AC)	2012: 0.596 (P ₉₅ : 0.630) 2014: 0.541 (P ₉₅ : 0.579) 2016: 0.543 (P ₉₅ : 0.580) 2018: 0.585 (P ₉₅ : 0.616) 2020: 0.570 (P ₉₅ : 0.603)	2012: 0.515 (P ₉₅ : 0.555) 2014: 0.609 (P ₉₅ : 0.646) 2016: 0.506 (P ₉₅ : 0.547) 2018: 0.535 (P ₉₅ : 0.571) 2020: 0.552 (P ₉₅ : 0.588)	
Turnover intention (TI)	2012: 0.619 (P ₉₅ : 0.662) 2014: 0.583 (P ₉₅ : 0.626) 2016: 0.559 (P ₉₅ : 0.600) 2018: 0.599 (P ₉₅ : 0.632) 2020: 0.539 (P ₉₅ : 0.576)	2012: 0.484 (P ₉₅ : 0.532) 2014: 0.472 (P ₉₅ : 0.521) 2016: 0.497 (P ₉₅ : 0.541) 2018: 0.452 (P ₉₅ : 0.496) 2020: 0.400 (P ₉₅ : 0.448)	2012: 0.514 (P ₉₅ : 0.561) 2014: 0.528 (P ₉₅ : 0.575) 2016: 0.501 (P ₉₅ : 0.543) 2018: 0.538 (P ₉₅ : 0.574) 2020: 0.574 (P ₉₅ : 0.610)

Values in parentheses report the 95th percentile of the percentile bootstrap distribution using 10,000 subsamples (i.e., P₉₅)



Table 8 Variance inflation factor (VIF) results of formative measurement models

Constructs	Indicators	Variance inflation factor (VIF) results				
		2012	2014	2016	2018	2020
Met expectations(ME)	PV1_diff	1.516	1.641	1.589	1.495	1.490
	PV2_diff	1.356	1.340	1.295	1.289	1.284
	PV3_diff	1.628	1.619	1.658	1.555	1.546
	PV4_diff	1.650	1.620	1.638	1.532	1.535
	PV5_diff	1.850	1.751	1.931	1.846	1.860
	PV6_diff	1.677	1.577	1.800	1.698	1.611
	PV7_diff	1.390	1.396	1.332	1.273	1.309
Motivational work characteristics (MWC)	feedback	1.170	1.119	1.117	1.150	1.120
	tasksign	1.090	1.058	1.049	1.051	1.054
	taskident	1.089	1.065	1.077	1.099	1.069

Table 9 Weights of formative measurement models

Constructs	Indicators	Results of the outer weights				
		2012	2014	2016	2018	2020
Met expectations(ME)	PV1_diff	0.269*	0.306*	0.242*	0.356*	0.410*
	PV2_diff	− 0.049	0.006	0.036	0.094*	0.056
	PV3_diff	0.387*	0.400*	0.318*	0.313*	0.249*
	PV4_diff	0.059	0.090	0.163*	0.034	− 0.022
	PV5_diff	0.219*	0.081	0.234*	0.280*	0.278*
	PV6_diff	0.200*	0.224*	0.163*	0.140*	0.240*
	PV7_diff	0.260*	0.250*	0.228*	0.185*	0.160*
Motivational work characteristics (MWC)	feedback	0.759*	0.689*	0.641*	0.706*	0.715*
	taskident	0.337*	0.360*	0.333*	0.333*	0.411*
	tasksign	0.225*	0.367*	0.453*	0.353*	0.272*

*Significant based on the 95% bias-corrected percentile bootstrap confidence intervals (10,000 bootstrap subsamples)

Table 10 Direct effects

	Direct effects	2012 estimate	2014 estimate	2016 estimate	2018 estimate	2020 estimate
		[95% CI]	[95% CI]	[95% CI]	[95% CI]	[95% CI]
<i>MWC</i> Motivational work characteristics, <i>ME</i> Met expectations, <i>LMX</i> Leader-membership exchange, <i>AC</i> Affective commitment, <i>JS</i> Job satisfaction, <i>TI</i> Turnover intention; CI=95% bias-corrected percentile bootstrap confidence intervals based on 10,000 bootstrap subsamples.	AC → TI	− 0.177 [− 0.234; − 0.120]	− 0.236 [− 0.292; − 0.178]	− 0.232 [− 0.287; − 0.173]	− 0.234 [− 0.283; − 0.184]	− 0.307 [− 0.354; − 0.258]
	JS → AC	0.379 [0.322; 0.438]	0.254 [0.199; 0.314]	0.321 [0.260; 0.383]	0.358 [0.309; 0.408]	0.334 [0.289; 0.383]
	JS → TI	− 0.441 [− 0.499; − 0.381]	− 0.403 [− 0.458; − 0.349]	− 0.385 [− 0.442; − 0.326]	− 0.414 [− 0.462; − 0.362]	− 0.311 [− 0.361; − 0.261]
	LMX → AC	0.162 [0.108; 0.217]	0.310 [0.253; 0.367]	0.177 [0.119; 0.236]	0.199 [0.152; 0.247]	0.229 [0.180; 0.278]
	LMX → JS	0.344 [0.288; 0.400]	0.299 [0.242; 0.358]	0.338 [0.290; 0.387]	0.277 [0.231; 0.326]	0.316 [0.270; 0.366]
	ME → AC	0.112 [0.053; 0.162]	0.102 [0.036; 0.155]	0.139 [0.081; 0.192]	0.118 [0.067; 0.164]	0.163 [0.115; 0.204]
	ME → JS	0.338 [0.282; 0.391]	0.347 [0.284; 0.404]	0.344 [0.286; 0.392]	0.345 [0.291; 0.392]	0.277 [0.222; 0.324]
	MWC → AC	0.091 [0.043; 0.135]	0.120 [0.070; 0.166]	0.098 [0.052; 0.141]	0.124 [0.085; 0.162]	0.075 [0.032; 0.116]
	MWC → JS	0.158 [0.116; 0.202]	0.144 [0.097; 0.192]	0.110 [0.062; 0.160]	0.153 [0.112; 0.193]	0.147 [0.106; 0.187]



Table 11 Total indirect effects

Indirect effects	2012 estimate [95% CI]	2014 estimate [95% CI]	2016 estimate [95% CI]	2018 estimate [95% CI]	2020 estimate [95% CI]
LMX → JS → AC	0.131 [0.104; 0.163]	0.076 [0.056; 0.103]	0.108 [0.083; 0.139]	0.099 [0.078; 0.125]	0.106 [0.085; 0.132]
ME → JS → AC	0.128 [0.101; 0.160]	0.088 [0.064; 0.117]	0.110 [0.084; 0.140]	0.124 [0.099; 0.150]	0.092 [0.071; 0.115]
MWC → JS → AC	0.060 [0.043; 0.081]	0.037 [0.023; 0.053]	0.035 [0.020; 0.055]	0.055 [0.039; 0.072]	0.049 [0.035; 0.065]
JS → AC → TI	− 0.067 [− 0.093; − 0.046]	− 0.060 [− 0.080; − 0.044]	− 0.074 [− 0.098; − 0.056]	− 0.084 [− 0.106; − 0.065]	− 0.102 [− 0.124; − 0.084]
LMX → JS → TI	− 0.152 [− 0.189; − 0.119]	− 0.121 [− 0.153; − 0.094]	− 0.130 [− 0.161; − 0.103]	− 0.115 [− 0.142; − 0.092]	− 0.098 [− 0.124; − 0.077]
LMX → AC → TI	− 0.029 [− 0.045; − 0.017]	− 0.073 [− 0.098; − 0.052]	− 0.041 [− 0.062; − 0.025]	− 0.047 [− 0.063; − 0.033]	− 0.070 [− 0.091; − 0.052]
ME → JS → TI	− 0.149 [− 0.181; − 0.120]	− 0.140 [− 0.173; − 0.107]	− 0.132 [− 0.163; − 0.102]	− 0.143 [− 0.173; − 0.113]	− 0.086 [− 0.108; − 0.065]
ME → AC → TI	− 0.020 [− 0.032; − 0.009]	− 0.024 [− 0.040; − 0.009]	− 0.032 [− 0.050; − 0.017]	− 0.028 [− 0.042; − 0.015]	− 0.050 [− 0.067; − 0.034]
MWC → JS → TI	− 0.070 [− 0.091; − 0.050]	− 0.058 [− 0.081; − 0.038]	− 0.042 [− 0.064; − 0.024]	− 0.063 [− 0.082; − 0.046]	− 0.046 [− 0.061; − 0.032]
MWC → AC → TI	− 0.016 [− 0.027; − 0.008]	− 0.028 [− 0.043; − 0.016]	− 0.023 [− 0.036; − 0.012]	− 0.029 [− 0.041; − 0.019]	− 0.023 [− 0.037; − 0.010]
LMX → JS → AC → TI	− 0.023 [− 0.033; − 0.016]	− 0.018 [− 0.026; − 0.013]	− 0.025 [− 0.035; − 0.018]	− 0.023 [− 0.031; − 0.017]	− 0.032 [− 0.042; − 0.025]
ME → JS → AC → TI	− 0.023 [− 0.033; − 0.015]	− 0.021 [− 0.029; − 0.015]	− 0.026 [− 0.035; − 0.018]	− 0.029 [− 0.038; − 0.021]	− 0.028 [− 0.037; − 0.021]
MWC → JS → AC → TI	− 0.011 [− 0.017; − 0.007]	− 0.009 [− 0.013; − 0.005]	− 0.008 [− 0.013; − 0.005]	− 0.013 [− 0.018; − 0.009]	− 0.015 [− 0.020; − 0.011]

MWC Motivational work characteristics, *ME* Met expectations, *LMX* Leader-membership exchange, *AC* Affective commitment, *JS* Job satisfaction, *TI* Turnover intention; CI=95% bias-corrected percentile bootstrap confidence intervals based on 10,000 bootstrap subsamples



Table 12 Total effects

Total effects	2012 estimate [95% CI]	2014 estimate [95% CI]	2016 estimate [95% CI]	2018 estimate [95% CI]	2020 estimate [95% CI]
JS → TI	− 0.508 [− 0.557; − 0.459]	− 0.463 [− 0.514; − 0.412]	− 0.459 [− 0.509; − 0.407]	− 0.498 [− 0.539; − 0.455]	− 0.413 [− 0.458; − 0.368]
LMX → AC	0.293 [0.241; 0.348]	0.386 [0.332; 0.442]	0.285 [0.232; 0.341]	0.299 [0.252; 0.347]	0.335 [0.288; 0.383]
LMX → TI	− 0.204 [− 0.241; − 0.170]	− 0.212 [− 0.248; − 0.179]	− 0.196 [− 0.227; − 0.167]	− 0.185 [− 0.215; − 0.158]	− 0.201 [− 0.231; − 0.173]
ME → AC	0.240 [0.183; 0.288]	0.190 [0.128; 0.241]	0.250 [0.194; 0.298]	0.242 [0.191; 0.286]	0.255 [0.208; 0.296]
ME → TI	− 0.192 [− 0.224; − 0.159]	− 0.185 [− 0.217; − 0.149]	− 0.190 [− 0.220; − 0.156]	− 0.200 [− 0.230; − 0.167]	− 0.164 [− 0.190; − 0.136]
MWC → AC	0.151 [0.102; 0.198]	0.156 [0.107; 0.203]	0.133 [0.086; 0.177]	0.179 [0.137; 0.217]	0.124 [0.080; 0.166]
MWC → TI	− 0.097 [− 0.121; − 0.072]	− 0.095 [− 0.120; − 0.069]	− 0.073 [− 0.098; − 0.049]	− 0.105 [− 0.128; − 0.083]	− 0.084 [− 0.105; − 0.062]

The direct effects reported in Table A7 have been excluded from this table.

MWC Motivational work characteristics, *ME* Met expectations, *LMX* Leader-membership exchange, *AC* Affective commitment, *JS* Job satisfaction, *TI* Turnover intention; CI=95% bias-corrected percentile bootstrap confidence intervals based on 10,000 bootstrap subsamples

Table 13 Cross-validated predictive ability test (CVPAT) results of the indicator average (IA) prediction benchmark

Construct	PLS- SEM loss	IA loss	Average loss difference	<i>t</i> -value	<i>p</i> -value
2012					
JS	1.851	3.210	− 1.359	12.764	0.000
AC	0.955	1.224	− 0.269	10.207	0.000
TI	1.651	1.855	− 0.203	8.850	0.000
Overall	1.381	1.778	− 0.397	13.796	0.000
2014					
JS	1.800	2.911	− 1.111	10.836	0.000
AC	0.844	1.146	− 0.303	11.397	0.000
TI	1.622	1.857	− 0.235	9.292	0.000
Overall	1.314	1.703	− 0.389	13.559	0.000
2016					
JS	1.703	2.786	− 1.083	10.857	0.000
AC	0.975	1.239	− 0.263	10.027	0.000
TI	1.560	1.832	− 0.272	11.131	0.000
Overall	1.330	1.714	− 0.384	14.099	0.000
2018					
JS	2.198	3.404	− 1.206	11.790	0.000
AC	0.933	1.225	− 0.293	11.960	0.000
TI	1.596	1.859	− 0.263	11.549	0.000
Overall	1.398	1.808	− 0.410	14.953	0.000
2020					
JS	2.008	2.898	− 0.890	10.877	0.000
AC	0.964	1.257	− 0.293	12.443	0.000
TI	1.513	1.692	− 0.179	9.549	0.000
Overall	1.349	1.678	− 0.330	13.813	0.000

AC Affective commitment, *JS* Job satisfaction, *TI* Turnover intention



Table 14 Cross-validated predictive ability test (CVPAT) results of the linear model (LM) prediction benchmark

	PLS-SEM loss	LM loss	Average loss difference	<i>t</i> -value	<i>p</i> -value
2012					
JS	1.851	1.851	− 0.000	0.018	0.986
AC	0.955	0.958	− 0.003	0.693	0.488
TI	1.651	1.640	0.012	1.134	0.257
Overall	1.381	1.378	0.004	0.753	0.451
2014					
JS	1.800	1.787	0.013	0.821	0.412
AC	0.844	0.845	− 0.002	0.385	0.700
TI	1.622	1.615	0.007	0.683	0.495
Overall	1.314	1.310	0.004	0.784	0.433
2016					
JS	1.703	1.688	0.015	1.073	0.283
AC	0.975	0.980	− 0.005	1.308	0.191
TI	1.560	1.548	0.012	1.048	0.295
Overall	1.330	1.325	0.005	0.921	0.357
2018					
JS	2.198	2.193	0.005	0.422	0.673
AC	0.933	0.924	0.009	1.722	0.085
TI	1.596	1.576	0.020	2.168	0.030
Overall	1.398	1.385	0.013	2.597	0.009
2020					
JS	2.008	1.991	0.016	1.224	0.221
AC	0.964	0.963	0.001	0.321	0.748
TI	1.513	1.505	0.009	1.143	0.253
Overall	1.349	1.342	0.007	1.596	0.111

AC Affective commitment, JS
Job satisfaction, TI Turnover
intention



Table 15 Permutation-based multigroup analysis (MGA) results of direct effects

Groups	Path	Direct effect time 1	Direct effect time 2	Diff. T1-T2	2.5%	97.5%	Permutation <i>p</i> value
T1: 2012 / T2: 2014	AC→TI	− 0.177	− 0.236	0.059	− 0.106	0.109	0.152
	JS→AC	0.379	0.254	0.125	− 0.108	0.109	0.002
	JS→TI	− 0.441	− 0.403	− 0.038	− 0.107	0.107	0.363
	LMX→AC	0.162	0.310	− 0.147	− 0.103	0.103	0.000*
	LMX→JS	0.344	0.299	0.045	− 0.109	0.105	0.274
	ME→AC	0.112	0.102	0.010	− 0.102	0.105	0.805
	ME→JS	0.338	0.347	− 0.008	− 0.106	0.108	0.838
	MWC→AC	0.091	0.120	− 0.029	− 0.091	0.086	0.392
	MWC→JS	0.158	0.144	0.014	− 0.083	0.084	0.665
T1: 2014 / T2: 2016	AC→TI	− 0.236	− 0.232	− 0.004	− 0.099	0.098	0.918
	JS→AC	0.254	0.321	− 0.067	− 0.112	0.112	0.125
	JS→TI	− 0.403	− 0.385	− 0.018	− 0.105	0.107	0.661
	LMX→AC	0.310	0.177	0.133	− 0.107	0.107	0.001(*)
	LMX→JS	0.299	0.338	− 0.039	− 0.099	0.096	0.309
	ME→AC	0.102	0.139	− 0.037	− 0.105	0.102	0.362
	ME→JS	0.347	0.344	0.003	− 0.103	0.102	0.936
	MWC→AC	0.120	0.098	0.022	− 0.085	0.089	0.506
	MWC→JS	0.144	0.110	0.034	− 0.090	0.087	0.334
T1: 2016 / T2: 2018	AC→TI	− 0.232	− 0.234	0.003	− 0.104	0.096	0.943
	JS→AC	0.321	0.358	− 0.037	− 0.101	0.105	0.352
	JS→TI	− 0.385	− 0.414	0.029	− 0.102	0.097	0.454
	LMX→AC	0.177	0.199	− 0.023	− 0.097	0.094	0.568
	LMX→JS	0.338	0.277	0.061	− 0.091	0.095	0.095
	ME→AC	0.139	0.118	0.021	− 0.096	0.096	0.562
	ME→JS	0.344	0.345	− 0.002	− 0.099	0.094	0.969
	MWC→AC	0.098	0.124	− 0.026	− 0.075	0.082	0.391
	MWC→JS	0.110	0.153	− 0.043	− 0.080	0.081	0.188
T1: 2018 / T2: 2020	AC→TI	− 0.234	− 0.307	0.073	− 0.092	0.092	0.042
	JS→AC	0.358	0.334	0.024	− 0.093	0.090	0.485
	JS→TI	− 0.414	− 0.311	− 0.103	− 0.093	0.096	0.005
	LMX→AC	0.199	0.229	− 0.030	− 0.087	0.088	0.384
	LMX→JS	0.277	0.316	− 0.039	− 0.090	0.086	0.262
	ME→AC	0.118	0.163	− 0.044	− 0.084	0.087	0.188
	ME→JS	0.345	0.277	0.069	− 0.094	0.092	0.057
	MWC→AC	0.124	0.075	0.049	− 0.076	0.075	0.089
	MWC→JS	0.153	0.147	0.006	− 0.077	0.072	0.847
T1: 2012 / T2: 2020	AC→TI	− 0.177	− 0.307	0.129	− 0.101	0.098	0.001(*)
	JS→AC	0.379	0.334	0.045	− 0.094	0.094	0.247
	JS→TI	− 0.441	− 0.311	− 0.130	− 0.102	0.102	0.001(*)
	LMX→AC	0.162	0.229	− 0.067	− 0.099	0.096	0.074
	LMX→JS	0.344	0.316	0.028	− 0.102	0.097	0.447
	ME→AC	0.112	0.163	− 0.050	− 0.092	0.092	0.152
	ME→JS	0.338	0.277	0.062	− 0.098	0.096	0.109
	MWC→AC	0.091	0.075	0.016	− 0.080	0.079	0.610
	MWC→JS	0.158	0.147	0.011	− 0.079	0.082	0.721

MWC Motivational work characteristics, *ME* Met expectations, *LMX* Leader-membership exchange, *AC* Affective commitment, *JS* Job satisfaction, *TI*=Turnover intention; * indicates significant differences between groups after applying the Bonferroni correction, which adjusted the significance level from 0.05 to 0.001; (*) indicates borderline *p* values of 0.001



Acknowledgements We used the statistical software SmartPLS (<https://www.smartpls.com>) in this research article. Christian M. Ringle acknowledges that he is a co-founder and co-developer of SmartPLS who has a financial interest in SmartPLS.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adler, S. J., L. Röseler, and M. K. Schöniger. 2023a. A toolbox to evaluate the trustworthiness of published findings. *Journal of Business Research* 167: 114189.
- Adler, S. J., P. N. Sharma, and L. Radomir. 2023b. Toward open science in PLS-SEM: Assessing the state of the art and future perspectives. *Journal of Business Research* 169: 114291.
- Altman, D. G., and J. M. Bland. 1995. Absence of evidence is not evidence of absence. *BMJ* 311(7003): 485.
- American Psychological Association. 2020. *Publication manual of the American Psychological Association*, 7th ed. Washington, DC: American Psychological Association.
- Bagozzi, R. P., and U. M. Dholakia. 2002. Intentional social action in virtual communities. *Journal of Interactive Marketing* 16(2): 2–21.
- Basco, R., J. F. Hair, C. M. Ringle, and M. Sarstedt. 2021. Advancing family business research through modeling nonlinear relationships: Comparing PLS-SEM and multiple regression. *Journal of Family Business Strategy* 13: 100457.
- Becker, J.-M., D. Proksch, and C. M. Ringle. 2022. Revisiting Gaussian copulas to handle endogenous regressors. *Journal of the Academy of Marketing Science* 50:46–66.
- Becker, J.-M., A. Rai, C. M. Ringle, and F. Völckner. 2013. Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly* 37(3): 665–694.
- Becker, J.-M., N. F. Richter, C. M. Ringle, and M. Sarstedt. 2026. Must-have, or maybe not? A sensitivity-based extension to necessary condition analysis. *Journal of Business Research* 206: 115920.
- Bell, E., B. Harley, and A. Bryman. 2022. *Business research methods*. Oxford: Oxford University Press.
- Biswas, D., and C. Szocs. 2019. The smell of healthy choices: Cross-modal sensory compensation effects of ambient scent on food purchases. *Journal of Marketing Research* 56(1): 123–141.
- Bollen, K. A., and P. J. Curran. 2006. *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.
- Cannon, J. P., and C. Homburg. 2001. Buyer–supplier relationships and customer firm costs. *Journal of Marketing* 65(1): 29–43.
- Chan, D. 1998. The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods* 1(4): 421–483.
- Cheah, J.-H., S. Amaro, and J. L. Roldán. 2023. Multigroup analysis of more than two groups in PLS-SEM: A review, illustration, and recommendations. *Journal of Business Research* 156: 113539.
- Cheah, J.-H., Sarstedt, M., Hair, J.F. and Ringle, C.M. 2026. Consistent partial least squares structural equation modeling using SmartPLS. *Structural Equation Modeling: A Multidisciplinary Journal*, forthcoming.
- Chin, W. W., and J. Dibbern. 2010. An introduction to a permutation based procedure for multi-group PLS analysis: Results of tests of differences on simulated data and a cross cultural analysis of the sourcing of information system services between Germany and the USA. In *Handbook of partial least squares: Concepts, methods and applications*, ed. V. Esposito Vinzi, W. W. Chin, J. Henseler, and H. Wang, 171–193. Berlin: Springer.
- Cohen, J. 1994. The earth is round ($p < .05$). *American Psychologist* 49(12): 997–1003.
- Collins, L. M. 2006. Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology* 57:505–528.
- Coltman, T., T. M. Devinney, D. F. Midgley, and S. Venaik. 2008. Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research* 61(12): 1250–1262.
- Curran, P. J., K. Obeidat, and D. Losardo. 2010. Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development* 11(2): 121–136.
- Dash, G., and J. Paul. 2021. CB-SEM vs PLS-SEM methods for research in social sciences and technology forecasting. *Technological Forecasting and Social Change* 173: 121092.
- Deer, L., S. J. Adler, H. Datta, N. Mizik, and M. Sarstedt. 2025. Toward open science in marketing research. *International Journal of Research in Marketing* 42(1): 212–233.
- Diamantopoulos, A., and N. Papadopoulos. 2010. Assessing the cross-national invariance of formative measures: Guidelines for international business researchers. *Journal of International Business Studies* 41(2): 360–370.
- Diamantopoulos, A., and J. A. Siguaw. 2000. *Introducing LISREL*. London: Sage.
- Diamantopoulos, A., and H. M. Winklhofer. 2001. Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research* 38(2): 269–277.
- Dul, J. 2016. Necessary condition analysis (NCA): Logic and methodology of “necessary but not sufficient” causality. *Organizational Research Methods* 19(1): 10–52.
- Ehrlich, J. S. 2024. *Die Entwicklung, Analyse und Replikation eines modifizierten Modells der Kündigungsabsicht*, vol. Vol. 66. Wiesbaden: Springer.
- Franke, G. R., and M. Sarstedt. 2019. Heuristics versus statistics in discriminant validity testing: A comparison of four procedures. *Internet Research* 29(3): 430–447.
- Franco, A., N. Malhotra, and G. Simonovits. 2014. Publication bias in the social sciences: Unlocking the file drawer. *Science* 345(6203): 1502–1505.
- French, B. F., and W. H. Finch. 2008. Multigroup confirmatory factor analysis: Locating the invariant referent sets. *Structural Equation Modeling: A Multidisciplinary Journal* 15(1): 96–113.
- Gelman, A., and J. Carlin. 2014. Beyond power calculations: Assessing Type S (sign) and Type M (magnitude) errors. *Perspectives on Psychological Science* 9(6): 641–651.
- George, J. M., and G. R. Jones. 2000. The role of time in theory and theory building. *Journal of Management* 26(4): 657–684.



- Gigerenzer, G., Krauss, S., Vitouch, O. 2004. The null ritual. *The Sage handbook of quantitative methodology for the social sciences*: 391–408.
- Girard, A., M. Lichters, M. Sarstedt, and D. Biswas. 2019. Short- and long-term effects of nonconsciously processed ambient scents in a servicescape: Findings from two field experiments. *Journal of Service Research* 22(4): 440–455.
- Greenland, S., S. J. Senn, K. J. Rothman, J. B. Carlin, C. Poole, S. N. Goodman, and D. G. Altman. 2016. Statistical tests, p values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology* 31(4): 337–350.
- Grimm, M., and R. Wagner. 2020. *The impact of missing values on PLS, ML and FIML model fit*. Karlsruhe: KIT Scientific Publishing.
- Grote, G., Staffelbach, B., Morf, M., Feierabend, A., Arnold, A., Pfrombeck, J. and Doden, W. 2017. Schweizer Human-Relations Barometer 2014 (Version 1.0.0) . FORS Datenservice.
- Grote, G., Tschopp, C., Doden, W., Staffelbach, B., Feierabend, A., Arnold, A. and Morf, M. 2013. Befragung von Angestellten zur Arbeits- und Beschäftigungssituation 2012 . FORS Datenservice.
- Guenther, P., M. Guenther, C. M. Ringle, G. Zaefarian, and S. Cartwright. 2023. Improving PLS-SEM use for business marketing research. *Industrial Marketing Management* 111:127–142.
- Guenther, P., M. Guenther, C. M. Ringle, G. Zaefarian, and S. Cartwright. 2025. PLS-SEM and reflective constructs: A response to recent criticism and a constructive path forward. *Industrial Marketing Management* 128:1–9.
- Hackman, J. R., and G. R. Oldham. 1976. Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance* 16(2): 250–279.
- Hahn, C., M. D. Johnson, A. Herrmann, and F. Huber. 2002. Capturing customer heterogeneity using a finite mixture PLS approach. *Schmalenbach Business Review* 54(3): 243–269.
- Hair, J. F., B. J. Babin, and N. Krey. 2017. Covariance-based structural equation modeling in the Journal of Advertising: Review and recommendations. *Journal of Advertising* 46(1): 163–177.
- Hair, J. F., G. T. M. Hult, C. M. Ringle, and M. Sarstedt. 2027. *A primer on partial least squares structural equation modeling (PLS-SEM)*, 4th ed. Thousand Oaks, CA: Sage.
- Hair, J. F., M. Sarstedt, L. Matthews, and C. M. Ringle. 2016. Identifying and treating unobserved heterogeneity with FIMIX-PLS: Part I – Method. *European Business Review* 28(1): 63–76.
- Hair, J. F., M. Sarstedt, C. M. Ringle, and S. P. Gudergan. 2024. *Advanced issues in partial least squares structural equation modeling*, 2nd ed. Thousand Oaks, CA: Sage.
- Hair, J. F., J. J. Risher, M. Sarstedt, and C. M. Ringle. 2019. When to use and how to report the results of PLS-SEM. *European Business Review* 31(1): 2–24.
- Hair, J. F., P. N. Sharma, W. W. Chin, M. Sarstedt, and C. M. Ringle. 2026. A multimethod SEM framework for analyzing models with latent variables. *Journal of Global Marketing*, forthcoming.
- Hamaker, E. L., R. M. Kuiper, and R. P. P. Grasman. 2015. A critique of the cross-lagged panel model. *Psychological Methods* 20(1): 102–116.
- Hawkins, A. T., and L. R. Samuels. 2021. Use of confidence intervals in interpreting nonstatistically significant results. *JAMA* 326(20): 2068–2069.
- Henseler, J., T. K. Dijkstra, M. Sarstedt, C. M. Ringle, A. Diamantopoulos, D. W. Straub, D. J. Ketchen, J. F. Hair, G. T. M. Hult, and R. J. Calantone. 2014. Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013). *Organizational Research Methods* 17(2): 182–209.
- Henseler, J., C. M. Ringle, and M. Sarstedt. 2015. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science* 43(1): 115–135.
- Henseler, J., C. M. Ringle, and M. Sarstedt. 2016. Testing measurement invariance of composites using partial least squares. *International Marketing Review* 33(3): 405–431.
- Hoenig, J. M., and D. M. Heisey. 2001. The abuse of power: The pervasive fallacy of power calculations for data analysis. *The American Statistician* 55(1): 19–24.
- Hom, P. W., T. W. Lee, J. D. Shaw, and J. P. Hausknecht. 2017. One hundred years of employee turnover theory and research. *Journal of Applied Psychology* 102(3): 530–545.
- Ioannidis, J. P. A. 2005. Why most published research findings are false. *PLoS Medicine* 2(8): e124.
- Johnston, M. P. 2014. Secondary data analysis: A method of which the time has come. *Qualitative and Quantitative Methods in Libraries* 3:619–626.
- Judd, C. M., and D. A. Kenny. 1981. Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review* 5(5): 602–619.
- Kerr, N. L. 1998. HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review* 2(3): 196–217.
- Kline, R. B. 2023. *Principles and practice of structural equation modeling*, 5th ed. New York, NY: Guilford.
- Knight, J. 2003. Null and void. *Nature* 422(6930): 554–555.
- Kozlov, M. 2024. So you got a null result. Will anyone publish it? *Nature* 631(8022): 728–730.
- Petrescu, M. and Krishen, A. S. 2025. The significance of nothing: a call to discussion. *Journal of Marketing Analytics* 13, 1033–1034.
- Lakens, D. 2017. Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science* 8(4): 355–362.
- Lakens, D., N. McLatchie, P. M. Isager, A. M. Scheel, and Z. Dienes. 2020. Improving inferences about null effects with Bayes factors and equivalence tests. *The Journals of Gerontology: Series B* 75(1): 45–57.
- Landis, R. S., L. R. James, C. E. Lance, C. A. Pierce, and S. G. Rogelberg. 2014. When is nothing something? Editorial for the null results special issue of Journal of Business and Psychology. *Journal of Business and Psychology* 29(2): 163–167.
- Legate, A. E., J. F. Hair, J. L. Chretien, and J. J. Risher. 2023. PLS-SEM: Prediction-oriented solutions for HRD researchers. *Human Resource Development Quarterly* 34(1): 91–109.
- Legate, A. E., C. M. Ringle, and J. F. Hair. 2024. PLS-SEM: A method demonstration in the R statistical environment. *Human Resource Development Quarterly* 35(4): 501–529.
- Lienggaard, B. D., J.-M. Becker, M. Bennedsen, P. Heiler, L. N. Taylor, and C. M. Ringle. 2025. Dealing with regression models' endogeneity by means of an adjusted estimator for the Gaussian copula approach. *Journal of the Academy of Marketing Science* 53:279–299.
- Lienggaard, B., P. Sharma, G. T. M. Hult, M. Jensen, M. Sarstedt, J. F. Hair, and C. M. Ringle. 2021. Prediction: Coveted, yet forsaken? Introducing a cross-validated predictive ability test in partial least squares path modeling. *Decision Sciences* 52(2): 362–392.
- Little, T. D. 2013. *Longitudinal structural equation modeling*. New York: Guilford.
- Lohmöller, J.-B. 1989. *Latent variable path modeling with partial least squares*. Heidelberg: Physica.
- Margalina, V.-M., Kreienbaum, C., Hair, J.F., Becker, J.-M. and Ringle, C.M. 2026. Multiple linear and logistic regression analysis: A SmartPLS 4 software tutorial. *Journal of Marketing Analytics*, forthcoming.
- Matthews, L. 2017. Applying multi-group analysis in PLS-SEM: A step-by-step process. In *Partial least squares structural equation modeling: Basic concepts, methodological issues and applications*, ed. R. Noonan and H. Latan, 219–243. Cham: Springer.



- Matthews, L., M. Sarstedt, J. F. Hair, and C. M. Ringle. 2016. Identifying and treating unobserved heterogeneity with FIMIX-PLS: Part II – A case study. *European Business Review* 28(2): 208–224.
- McShane, B. B., D. Gal, A. Gelman, C. Robert, and J. L. Tackett. 2019. Abandon statistical significance. *The American Statistician* 73(sup1): 235–245.
- Meehl, P. E. 1978. Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology* 46(4): 806–834.
- Meehl, P. E. 1990. Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports* 66(1): 195–244.
- Meehl, P. E. 2004. Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Applied and Preventive Psychology* 11(1): 1–7.
- Melumad, S., J. J. Inman, and M. T. Pham. 2019. Selectively emotional: How smartphone use changes user-generated content. *Journal of Marketing Research* 56(2): 259–275.
- Memon, M. A., J.-H. Cheah, T. Ramayah, H. Ting, F. Chuah, and T. H. Cham. 2019. Moderation analysis: Issues and guidelines. *Journal of Applied Structural Equation Modeling* 3(1): i–xi.
- Memon, M. A., R. Thurasamy, J.-H. Cheah, H. Ting, C. Kreienbaum, and C. M. Ringle. 2026. Moderated mediation analysis using SmartPLS: A tutorial for business researchers. *Journal of Applied Structural Equation Modeling* 10(1): 1–28.
- Meredith, W. 1993. Measurement invariance, factor analysis, and factorial invariance. *Psychometrika* 58(4): 525–543.
- Merkle, A. C. 2025. Item level correction: Detecting, removing, and reporting common methods variance. *Journal of Marketing Analytics* 13(2): 405–423.
- Meyer, J. P., and N. J. Allen. 1991. A three-component conceptualization of organizational commitment. *Human Resource Management Review* 1(1): 61–89.
- Meyer, J. P., N. J. Allen, and C. A. Smith. 1993. Commitment to organizations and occupations: Extension and test of a three-component conceptualization. *Journal of Applied Psychology* 78(4): 538–551.
- Mobley, W. H. 1977. Intermediate linkages in the relationship between job satisfaction and employee turnover. *Journal of Applied Psychology* 62(2): 237–240.
- Morgeson, F. P., and S. E. Humphrey. 2006. The work design questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology* 91(6): 1321–1339.
- Mowday, R. T., R. M. Steers, and L. W. Porter. 1979. The measurement of organizational commitment. *Journal of Vocational Behavior* 14(2): 224–247.
- Nath, A., S. Schimmelpfennig, C. Kreienbaum, and U. Konradt. 2025. Move to improve: Meta-analysis of workplace physical activity interventions. *Journal of Occupational Health Psychology* 30(3): 176–198.
- Nickerson, R. S. 2000. Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods* 5(2): 241–301.
- Nosek, B. A., G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. L. Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. Vandenberg, S. Vazire, E. J. Wagenmakers, R. Wilson, and T. Yarkoni. 2015. Promoting an open research culture. *Science* 348(6242): 1422–1425.
- Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *Science* 349(6251): aac4716.
- Pfrombeck, J., A. Feierabend, L. Schärer, A. Kornblum, G. Grote, and B. Staffelbach. 2020. *Schweizer Human Relations Barometer 2020: Digitalisierung und Generationen*. Distributed by SWISS-Subase: ETH Zürich and Universitäten Luzern und Zürich.
- Pfrombeck, J., Schärer, L., Feierabend, A., Roth, M., Grote, G. and Staffelbach, B. 2018. Schweizer Human-Relations-Barometer 2018: Integration und Diskriminierung. FORS Datenservice.
- Ployhart, R. E., and R. J. Vandenberg. 2010. Longitudinal research: The theory, design, and analysis of change. *Journal of Management* 36(1): 94–120.
- Podsakoff, N. P., J. A. LePine, and M. A. LePine. 2007. Differential challenge stressor–hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: A meta-analysis. *Journal of Applied Psychology* 92(2): 438–454.
- Raeder, S., A. Wittekind, A. Inauen, and G. Grote. 2009. Testing a psychological contract measure in a Swiss employment context. *Swiss Journal of Psychology* 68(4): 177–188.
- Rigdon, E. E. 1998. Structural equation modeling. In *Modern methods for business research*, ed. G. A. Marcoulides, 251–294. Mahwah, NJ: Lawrence Erlbaum Associates.
- Rindfleisch, A., A. J. Malter, S. Ganesan, and C. Moorman. 2008. Cross-sectional versus longitudinal survey research: Concepts, findings, and guidelines. *Journal of Marketing Research* 45(3): 261–279.
- Ringle, C. M., and M. Sarstedt. 2016. Gain more insight from your PLS-SEM results: The importance-performance map analysis. *Industrial Management & Data Systems* 116(9): 1865–1886.
- Ringle, C. M., M. Sarstedt, R. Mitchell, and S. P. Gudergan. 2020. Partial least squares structural equation modeling in HRM research. *The International Journal of Human Resource Management* 31(12): 1617–1643.
- Ringle, C. M., M. Sarstedt, N. Sinkovics, and R. R. Sinkovics. 2023. A perspective on using partial least squares structural equation modelling in data articles. *Data in Brief* 48: 109074.
- Ringle, C. M., Wende, S. and Becker, J.-M. 2024. SmartPLS 4. Bönningstedt: SmartPLS. Available at: <https://www.smartpls.com>.
- Robinson, S. L., and E. Wolfe Morrison. 2000. The development of psychological contract breach and violation: A longitudinal study. *Journal of Organizational Behavior* 21(5): 525–546.
- Roemer, E. 2016. A tutorial on the use of PLS path modeling in longitudinal studies. *Industrial Management & Data Systems* 116(9): 1901–1921.
- Rosenthal, R. 1979. The file drawer problem and tolerance for null results. *Psychological Bulletin* 86(3): 638–641.
- Sarstedt, M. 2026. Metrological uncertainty and the value of informative nulls in marketing research. *Journal of Marketing Analytics*, forthcoming.
- Sarstedt, M., S. J. Adler, C. M. Ringle, G. Cho, A. Diamantopoulos, and H. Hwang. 2024. Same model, same data, but different outcomes: Evaluating the impact of method choices in structural equation modeling. *Journal of Product Innovation Management* 41(6): 1100–1117.
- Sarstedt, M., J. F. Hair, C. Nitzl, C. M. Ringle, and M. C. Howard. 2020a. Beyond a tandem analysis of SEM and PROCESS: Use of PLS-SEM for mediation analyses! *International Journal of Market Research* 62(3): 288–299.
- Sarstedt, M., C. M. Ringle, J.-H. Cheah, H. Ting, O. I. Moisesescu, and L. Radomir. 2020b. Structural model robustness checks in PLS-SEM. *Tourism Economics* 26(4): 531–554.
- Sarstedt, M., J. F. Hair, M. Pick, B. D. Lienggaard, L. Radomir, and C. M. Ringle. 2022a. Progress in partial least squares structural equation modeling use in marketing research in the last decade. *Psychology & Marketing* 39(5): 1035–1064.
- Sarstedt, M., L. Radomir, O. I. Moisesescu, and C. M. Ringle. 2022b. Latent class analysis in PLS-SEM: A review and



- recommendations for future applications. *Journal of Business Research* 138:398–407.
- Sarstedt, M., C. M. Ringle, and J. F. Hair. 2017. Treating unobserved heterogeneity in PLS-SEM: A multi-method approach. In *Partial least squares structural equation modeling: Basic concepts, methodological issues and applications*, ed. R. Noonan and H. Latan, 197–217. Cham: Springer.
- Sarstedt, M., C. M. Ringle, and J. F. Hair. 2025. Partial least squares structural equation modeling. In *Handbook of market research*, 1–56. Cham: Springer.
- Sarstedt, M., J. Henseler, and C. M. Ringle. 2011. Multigroup analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. In *Advances in international marketing*, ed. M. Sarstedt, M. Schwaiger, and C. R. Taylor, 195–218. Emerald Group Publishing Limited.
- Satorra, A., and P. M. Bentler. 2010. Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika* 75(2): 243–248.
- Schyns, B. 2002. Überprüfung einer deutschsprachigen Skala zum Leader-Member-Exchange-Ansatz. *Zeitschrift Für Differentielle und Diagnostische Psychologie* 23(2): 235–245.
- Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Sharma, P. N., B. D. Liengard, J. F. Hair, M. Sarstedt, and C. M. Ringle. 2023. Predictive model assessment and selection in composite-based modeling using PLS-SEM: Extensions and guidelines for using CVPAT. *European Journal of Marketing* 57(6): 1662–1677.
- Sharma, P. N., M. Sarstedt, C. M. Ringle, J.-H. Cheah, A. Herfurth, and J. F. Hair. 2024. A framework for enhancing the replicability of behavioral MIS research using prediction oriented techniques. *International Journal of Information Management* 78: 102805.
- Shmueli, G., M. Sarstedt, J. F. Hair, J.-H. Cheah, H. Ting, S. Vaithilingam, and C. M. Ringle. 2019. Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing* 53(11): 2322–2347.
- Simmons, J. P., L. D. Nelson, and U. Simonsohn. 2011. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science* 22(11): 1359–1366.
- Singer, J. D., and J. B. Willett. 2003. *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford: Oxford University Press.
- Staffelbach, B., Grote, G., Feierabend, A., Humm, J., Morf, M., Arnold, A., Pfrombeck, J. and Doden, W. 2022. Schweizer Human-Relations-Barometer 2016 (Version 1.0.0) . FORS Datenservice.
- Steel, R. P., and J. W. Lounsbury. 2009. Turnover process models: Review and synthesis of a conceptual literature. *Human Resource Management Review* 19(4): 271–282.
- Tehseen, S., T. Ramayah, and S. Sajilan. 2017. Testing and controlling for common method variance: A review of available methods. *Journal of Management Sciences* 4(2): 146–175.
- Vaithilingam, S., C. S. Ong, O. I. Moisescu, and M. S. Nair. 2024. Robustness checks in PLS-SEM: A review of recent practices and recommendations for future applications in business research. *Journal of Business Research* 173: 114465.
- Vandenberg, R. J. 2002. Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods* 5(2): 139–158.
- Vandenberg, R. J., and D. M. Grelle. 2010. Alternative model specifications in structural equation modeling: Facts, fictions, and truth. In *Statistical and methodological myths and urban legends: Doctrine, verity and fable in organizational and social sciences*, ed. C. E. Lance and R. J. Vandenberg, 165–191. New York, NY: Routledge.
- Vandenberg, R. J., and C. E. Lance. 2000. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods* 3(1): 4–70.
- Wang, H., S. Lu, and Y. Liu. 2022. Missing data imputation in PLS-SEM. *Quality & Quantity* 56:4777–4795.
- Wasserstein, R. L., A. L. Schirm, and N. A. Lazar. 2019. Moving to a world beyond “ $p < 0.05$.” *The American Statistician* 73(sup1): 1–19.
- Wold, H. 1982. Soft modelling: The basic design and some extensions. In *Systems under indirect observation: Causality, structure, prediction*, vol. 2, ed. K. G. Jöreskog and H. Wold, 1–54. Amsterdam: North-Holland.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Janna S. Ehrlich is a Professor of Business Psychology at Stralsund University of Applied Sciences (Germany). Her research focuses on organizational behavior, employee attitudes, and turnover intentions, with a particular emphasis on longitudinal modeling and structural equation modeling (SEM). Trained in business psychology, she works at the intersection of theory-driven research and applied organizational contexts. Her research has been presented at international conferences, including the European Association of Work and Organizational Psychology (EAWOP) and the International PLS-SEM Conference. By combining rigorous quantitative methods with practical relevance, her work contributes to a deeper understanding of employee behavior and evidence-based management in organizations.

Charlotte Kreienbaum is a research fellow at the Institute of Management and Decision Sciences at the Hamburg University of Technology (Germany) and a rising Ph.D. student whose research interests lie in organizational behavior and business analytics. Trained in psychology, she conducts research at the intersection of human behavior and data-driven decision making, with a particular interest in how organizations can optimize performance, well-being, and strategic outcomes. By combining quantitative methods with a behavioral science foundation, her work advances evidence-based insights for decision making in complex organizational and societal systems. Her scholarly contributions include publications in the *Journal of Occupational Health Psychology*, *Journal of Marketing Analytics*, and *Journal of Applied Structural Equation Modeling*.

Christian M. Ringle is a Chaired Professor and the Director of the Institute of Management and Decision Sciences at the Hamburg University of Technology (Germany) and an Adjunct Professor at the James Cook University (Australia). His research focuses on management, marketing, technology and innovation management, method development, business analytics, machine learning, artificial intelligence, and the application of business research methods to decision making. His contributions have been published in journals such as *Decision Sciences*, *European Journal of Marketing*, *International Journal of Research in Marketing*, *Information Systems Research*, *Journal of Business Research*, *Journal of Service Research*, *Journal of the Academy of Marketing Science*, *Long Range Planning*, *MIS Quarterly*, and *Organizational Research Methods*. Since 2018, Christian has been included in the Clarivate Analytics' list of Highly Cited Researchers (HCR). He is a co-developer and co-founder of the statistical software SmartPLS (<https://www.smartpls.com>). More info: <https://www.tuhh.de/mds/team/prof-dr-c-m-ringle>.

