Three Issues in Multilevel Research

Vicente González-Romá

Universitat de València (Spain)

Abstract. In this article, three important issues in organizational multilevel research are discussed and clarified, namely: (a) The interpretation of "cross-level direct effects" in theoretical and research multilevel models, (b) the specification of the emergence processes involved in higher-level constructs, and (c) the sample size recommendations for using multilevel statistical methods. By doing so, this article hopes to contribute to the improvement of organizational multilevel research.

Organizational phenomena often involve relationships, variables, and factors that reside at different levels[1] within organizations; that is, they have a multilevel nature. For instance, a researcher may think that: (a) Team support climate (i.e., shared perceptions of supportive relationships within the team) has a positive impact on employees' well-being at work, and (b) the negative relationship between employees' job burnout and well-being varies depending on team support climate, so that when team support climate is high, the individual-level relationship between job burnout and well-being is weakened. In this example, we see a *team*-level property (support climate) that hypothetically has an influence on an *individual*-level variable (employee well-being) and an *individual* relationship (employee job burnout → employee well-being). In order to fully understand these types of phenomena, we have to adopt a multilevel perspective.

However, in the last century, our discipline was dominated by a single-level approach (mainly, the individual level; Hitt, Beamish, Jackson, & Mathieu, 2007). Undoubtedly, this approach yielded great scientific advancement. However, it cannot explain the complexities of many organizational phenomena in which predictors, intervening variables, and outcomes reside at different levels (Kozlowski & Klein, 2000). In the last two decades of the 20th century, the multilevel paradigm in work and organizational psychology was introduced (House, Rousseau, & Thomas-Hunt, 1995; Klein, Dansereau, & Hall, 1994; Kozlowski & Klein, 2000; Rousseau, 1985). This paradigm posits that the properties of a specific entity (e.g., employees) are related to the properties of other entities that reside at different levels (e.g., work teams). In order to estimate these multilevel relationships accurately, new and appropriate statistical methods were needed. At the same time that the multilevel paradigm in work and organizational psychology was initiated, some statisticians started to develop the multilevel methods and software tools needed to examine multilevel relationships (Burstein, Linn, & Capell, 1978; de Leeuw & Kreft, 1986; Goldstein, 1986; Muthén, 1989). The confluence of these two substantive and methodological streams yielded a steady increase in the number of multilevel studies in our discipline from the turn of the century on (see González-Romá & Hernández, 2017). This increase in the number of published multilevel studies is good news because it shows that researchers are becoming aware that multilevel organizational phenomena need appropriate models, concepts, and methods. However, it has also raised some issues that require further elaboration and clarification. Three of these issues are the following: (a) The interpretation of "cross-level direct effects" in theoretical and research multilevel models, (b) the specification of the emergence processes involved in higher-level constructs, and (c) the sample size recommendations for using multilevel statistical methods. The goal in this article is to contribute to improving our understanding and use of multilevel models, concepts, and methods, by discussing and clarifying the three issues mentioned above.

[1]A level denotes a specific position in a series of nested arrangements (e.g., employees are nested in teams, teams in departments, and departments in organizations). Employees, teams, departments, and organizations are some of the levels that can be considered within organizations.

The interpretation of cross-level direct effects in theoretical and research multilevel models

Although this issue has been discussed previously in the literature (LoPilato & Vandenberg, 2015; see also Preacher, Zyphur, & Zhang, 2010), it continues to be a source of misunderstanding. For illustrative purposes, suppose that the imaginary researcher we introduced in the first paragraph estimates the individual-level relationship between employee job burnout and employee well-being in each of the work teams making up his/her sample[2]. [Remember that s/he does so based on the idea that this relationship may vary across work teams depending on a team characteristic (support climate)]. Thus, s/he estimates a simple regression model in each team: $Y = a + b X + e$, where Y is the outcome variable (well-being), X is the predictor variable (job burnout), *a* is the regression intercept, *b* is the regression coefficient or slope, and *e* is the residual term. Using a multilevel notation and taking into account the nested structure of the data, the estimated regression model can be expressed as follows:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + r_{ij} \qquad (1)$$

where Y_{ij} is the score on the outcome variable of subject *i* from team *j*, X_{ij} is the score on the predictor variable of subject *i* from team *j*, β_{0j} is the regression intercept estimated in each team (*j*), β_{1j} is the regression coefficient (slope) estimated in each team (*j*), and r_{ij} is the residual term of the regression equation in each team (*j*). Note that in this equation the outcome and the predictor variables are individual-level variables (in the example case, job burnout and well-being, respectively). Thus, equation 1 is an individual-level equation.

Suppose that our researcher represents the regression lines obtained in each team in his/her sample and obtains a graph similar to the one represented in Figure 1 (in which, for the sake of clarity, only the regression lines of five teams are displayed).

Figure 1 points out that different teams show quite different regression lines. This means that there are differences across teams in the regression intercept (β_{0j}, the point at which the regression lines cross the Y-axis) and the regression slope (β_{1j}). In fact, the regression intercept varies between 2.1 and 4.9, and the slope between −.72 and 0. Because our researcher is interested in examining the sources of these variabilities in the regression intercepts and slopes across the *j* teams of his/her sample, and based on his/her idea that this variance may depend on work teams' support climate (C_j), s/he writes the following simple regression equations of the $Y = a + b X + e$ form:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} C_j + U_{0j} \qquad (2)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} C_j + U_{1j} \qquad (3)$$

Note that in these equations, the outcome variables (β_{0j} and β_{1j}) are two team characteristics (teams' intercepts and slopes, respectively), and the predictor variable is also a team characteristic (support climate, C_j). Thus, equations 2 and 3 are team-level equations. Moreover, for the goals of this section, suffice it to say that γ_{00} and γ_{10} are two regression intercepts, γ_{01} and γ_{11} are two regression coefficients (slopes) that estimate the relationship between C_j, on the one hand, and β_{0j} and β_{1j}, on the other, and U_{0j} and U_{1j} are the corresponding residual terms.

Equations 1, 2, and 3 form a multilevel model in which different relationships are specified at different levels of analysis (the individual level: Equation 1, and the team level: Equations 2 and 3). Frequently, the terms β_{0j} and β_{1j} in Equation 1 are replaced with the corresponding right-hand parts of Equations 2 and 3, respectively. After some simple algebra and a rearrangement of terms, the following combined equation is obtained:

$$Y_{ij} = \gamma_{00} + \gamma_{01} C_j + \gamma_{10} X_{ij} + \gamma_{11}(C_j X_{ij}) + (U_{0j} + U_{1j} X_{ij} + r_{ij})$$
$$(4)$$

The regression coefficient γ_{01} is said to represent a so-called "cross-level direct effect", whereas the regression coefficient γ_{11} estimates a so-called "cross-level interaction effect" or "cross-level moderation". These two effects are usually represented as they appear in Figure 2. In my opinion, by focusing on equation 4 and the typical graphical representation of the cross-level direct effect, researchers have frequently interpreted the regression coefficient γ_{01} as estimating the relationship between a team-level predictor (e.g., team support climate, C_j) and an individual-level outcome (e.g., well-being, Y_{ij}) [see LoPilato and Vandenberg (2015) for examples]. This interpretation is inaccurate for several reasons.

Conceptually speaking, a team-level property such as support climate (C_j) cannot explain within-team variability in employees' scores on an outcome variable such as well-being (Y_{ij}) because any team-level property is constant within a specific team (Preacher et al., 2010). To illustrate this, in Figure 3 we display some cases from a multilevel data file that includes the variables used in our example. Within each team, the values for team support climate are the same for all the employees who belong to the same team (4.67 for Team 1, 4.33 for Team 2, 4.25 for Team 3). Team support climate is constant within teams. Because it is a team-level variable, there can be no within-team differences in this variable. Therefore, it cannot explain the differences in well-being observed within teams.

---

[2]For the sake of clarity, we will focus on an example involving two levels: employees (Level 1) and work teams (Level 2).
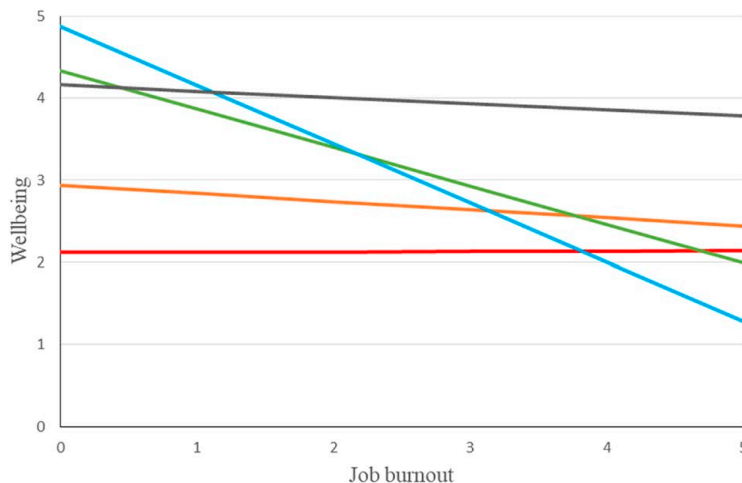
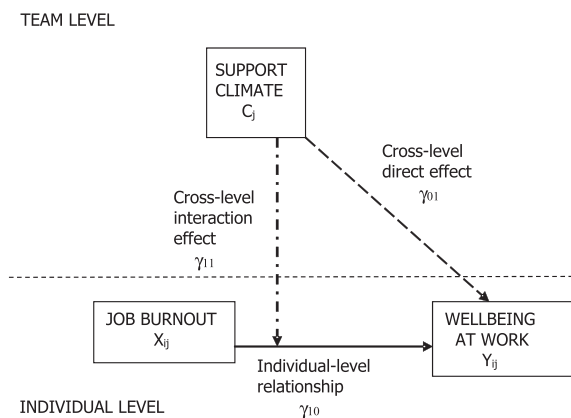**Figure 1.** Regression lines obtained after regressing well-being at work on job burnout in five work teams.

Figure 2. Typical representation of cross-level direct and interaction effects.

Team Code	Employee Code	Team Support Climate	Employee Job Burnout	Employee Well-being at Work
1	11	4.67	1.6	3.68
1	12	4.67	2.8	3.24
1	13	4.67	1.6	3.64
1	14	4.67	2.4	2.56
2	21	4.33	1.6	3.2
2	22	4.33	3.2	3.36
2	23	4.33	1.8	3.76
2	24	4.33	2.8	3.88
3	31	4.25	3.4	2.17
3	32	4.25	3	2.17
3	33	4.25	2.8	3.5
3	34	4.25	3.4	3.32

Figure 3. First cases from a multilevel data file.

Mathematically, Equation 2 tells us that γ_{01} estimates the relationship between two team variables: team support climate (C_j) and the team intercept (β_{0j}), rather than between C_j and employee well-being (Y_{ij}). Thus, it makes sense to clarify the meaning of β_{0j}. In a simple regression model ($Y = a + b\,X + e$), the value for the intercept is yielded by $a = \bar{Y} - b\,\bar{X}$. Applying this expression to equation 1, we can write the following:

$$\beta_{0j} = \bar{Y}_j - \beta_{1j}\,\bar{X}_j \tag{5}$$

Thus, in our example, β_{0j} can be interpreted as a team mean in well-being ($\bar{Y}_j$), once the effect of team average job burnout ($\beta_{1j}\,\bar{X}_j$) has been controlled for; that is, β_{0j} is an adjusted team mean in well-being (Enders & Tofighi, 2007; LoPilato & Vandenberg, 2015). If we replace β_{0j} in Equation 2 with the right-hand part of equation 5, we get the following expression (LoPilato & Vandenberg, 2015):

$$(\bar{Y}_j - \beta_{1j}\bar{X}_j) = \gamma_{00} + \gamma_{01}C_j + U_{0j} \tag{6}$$

Therefore, the coefficient that estimates the so-called "cross-level direct effect" (γ_{01}) provides an estimation of the relationship between the involved team-level predictor (in our example, team support climate, C_j) and an adjusted team mean in the outcome variable (in our example, well-being). It does not provide an estimation of the relationship between a team-level predictor (e.g., team support climate, C_j) and an individual-level outcome (e.g., well-being, Y_{ij})[3].

In order to avoid interpreting γ_{01} as estimating the "effect" of a team-level variable on an individual-level outcome, several steps can be taken. First, focus should be placed on equations 2, and 6, rather than on equation 4. Equations 2 and 6 make explicit that γ_{01} estimates the relationship between two team-level variables (a team-level predictor and a

---

[3]γ_{01} could be interpreted in this way only in a very extreme situation: when within-team agreement in the outcome variable is extremely high. See LoPilato and Vandenberg (2015) for a detailed discussion.

team intercept that, in this case, can be interpreted as an adjusted mean in the outcome variable).

Second, we should consider alternative ways of graphically representing the "cross-level direct effect". In two-level designs similar to the one used in our example (e.g., individuals nested within teams), the variance in individual-level variables can be partitioned into two orthogonal components: the Between-team component and the Within-team component[4] (Preacher et al., 2010). Because team-level variables can only vary across teams, they only have Between-team components of variance. γ_{01} estimates the relationship between two components: the team-level predictor (e.g., C_j) and the Between component of the individual-level outcome involved. It is a "Between effect" (Preacher et al., 2010). Thus, an alternative way of representing the "cross-level direct effect" is shown in Figure 4 (see LoPilato & Vandenberg, 2015, p. 295). Note that in this representation, there is no arrow going from the team-level predictor (C_j) to the individual-level outcome (Y_{ij}).

Third, as mentioned above, γ_{01} is a "Between effect". Therefore, we should consider replacing its typical label of "cross-level direct effect" with the label of "Between-direct effect". The "cross-level" term in the typical label suggests that the relationship in question runs across levels, from a higher-level variable to a lower-level one. This may have contributed to misinterpreting the meaning of γ_{01}.
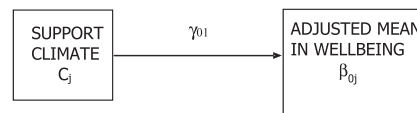
Fourth, we should formulate hypotheses and interpret the associated results more accurately. Using our example, a typical "cross-level direct effect" hypothesis would read as follows: "Team support climate is positively related to employee well-being, after employee job burnout in controlled for". However, a more precise hypothesis would be: "Team support climate is positively related to the teams' adjusted means in well-being, once the effect of team average job burnout is controlled for". The interpretation of the obtained results should be congruent with the hypothesis formulated and the meaning of γ_{01} discussed above. For instance, an estimate for γ_{01} of .21 ($p < .01$) would mean that there is a positive relationship between team support climate and teams' adjusted means on well-being. We could also say that there is a positive relationship between team support climate and the Between-component of well-being.

The specification of the emergence processes involved in higher-level constructs

Multilevel models in Work and Organizational Psychology frequently involve higher-level (e.g., team)
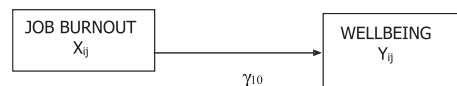

Figure 4. An alternative way of representing the so-called "cross-level direct effect".

constructs that have their origin in lower-level (e.g., individual) properties. For instance, in our example, team support climate (i.e., shared team members' perceptions about supportive relationships within the team) has its origins in team members' individual perceptions of support in the team. To fully understand the nature of higher –level constructs, it is of utmost importance to explain the processes through "which lower-level properties emerge to form collective phenomena" (Kozlowski & Klein, 2000, p. 15), that is, to explain the emergence processes involved. For instance, how do employees' team climate perceptions become shared to form team climate? In my experience as a journal editor, associate editor, and reviewer, I have observed that this theoretical explanation is frequently disregarded in research manuscripts. When authors work with shared higher-level constructs (such as team climate), it often seems that we are only concerned about showing that there is empirical evidence to justify aggregation of individual scores to obtain an indicator of the shared construct. Then, authors generally provide evidence about inter-rater agreement (usually by means of the r_{wg} index; James, Demaree, & Wolf, 1984), the percentage of variance that resides at the higher level (e.g., team), by reporting the corresponding Intraclass Correlation Coefficient (1), ICC(1), and the reliability of the team means by reporting the ICC (2). Of course, showing this evidence is operationally important, but understanding the nature of higher-level constructs and the processes involved in their emergence from lower-level properties is theoretically crucial. A brief discussion of general types of emergence processes with some specific examples can be useful to clarify this issue (for a more detailed discussion see Kozlowski & Klein, 2000).

Kozlowski and Klein (2000) describe two general types of emergence processes: Composition and compilation. In composition emergence, the type and amount

---

[4]This is demonstrated by the Intraclass Correlation Coefficient (ICC). Suppose that well-being showed an ICC of .22. This means that 22% of the variance in wellbeing is due to differences in teams' averaged scores in well-being (between-team differences), whereas 78% is due to individual differences within teams.

of the lower-level property involved is similar for all the members of the work unit. For instance, in the case of team support climate, the higher-level construct is formed from the same type of individual property (team members' perception of support, i.e., psychological climate), and all the team members perceive similar (shared) amounts of support. Composition processes of emergence explain how convergence, sharing, and within-unit agreement develop to yield a shared unit property (Kozlowski & Klein, 2000). For instance, organizational climate theory and research has identified social interaction among team members as a key process for team climate emergence (Ashforth, 1985; González-Romá & Peiró, 2014; González-Romá, Peiró, & Tordera, 2002; Rentsch, 1990). Through continued social interactions, team members communicate, discuss, and shape their perceptions of the team. These ongoing social interactions yield a shared perception of the team that is socially constructed (Ashforth, 1985). Other examples of composition processes of emergence can be found in the literature on collective mood (see Barsade & Knight, 2015; Kelly & Barsade, 2001).

In compilation emergence, either the amount or type of the lower-level property is different, "or both the amount and type are different" (Kozlowski & Klein, 2000, p. 62). To exemplify these cases, let us think first about the performance of a basketball team. To reach high performance, the basketball team needs to have a varied pool of individual skills (i.e., distinct lower-level properties). For instance, the point guard contributes to team performance by interpreting the dynamic development of the game and making strategy decisions. Likewise, the center contributes in another way by enacting different skills, such as occupying a good position under the basket and blocking the opponent's throws. Moreover, the players can have different or similar amounts of their respective skills.

The higher-level construct of climate uniformity (González-Romá & Hernández, 2014) exemplifies a case in which only the amount of the lower-level property is different. Climate uniformity refers to the pattern of climate perceptions within the team. It represents differences in within-unit (dis)agreement in climate perceptions across work units. González-Romá and Hernández (2014) examined three different patterns of climate perceptions of organizational support: Uniform, in which there is a single grouping of unit members' climate perceptions; strong dissimilarity, in which distinct subgroupings within the unit are observed, located in different zones (e.g., low vs high scores) of the underlying climate facet, as occurs with polarized subgroups; and weak dissimilarity patterns, in which no more than one subgrouping of climate perceptions is found, and the members excluded from this subgroup do not form a coherent cluster. Here, the type of individual property involved is the same for all the unit members: perceptions of organizational support. However, the amount of organizational support perceived by team members can be different, which can give rise to different unit climate patterns.

Compilation processes of emergence foster variability and configuration, and they explain how different types or/and amounts of lower-level properties combine to yield higher-order configural properties (Kozlowski & Klein, 2000). In the case of climate uniformity, González-Romá and Hernández (2014) suggest that leader–member exchange (LMX) quality, demographic diversity, and organizational socialization may explain the emergence of climate uniformity. For instance, in the case of LMX, they posit that "team leaders who relate differently with distinct subgroups of team members, or have differentiated relationships with specific individual team members, may contribute to fostering nonuniform team climate patterns [strong and weak dissimilarity patterns, respectively]" (pp. 1054–1055). These differentiated relationships may produce subgroupings of team members that have different climate perceptions about the team.

As Kozlowski and Klein (2000) highlight, when a higher-level construct that has its origin in lower-level properties is used in a multilevel model, the processes that explain how the corresponding lower-level properties emerge to form a higher-level construct must be explained. Therefore, I recommend that when authors refer to a higher-level construct in the Introduction to their manuscripts, they devote enough space to this explanation about the emergent processes involved. By doing so, they will contribute to improving our understanding of the nature of higher-level constructs.

Sample size recommendations for using multilevel statistical methods

An important issue to clarify when designing an empirical research study is to establish the size of the study sample. This issue is even more relevant in multilevel studies in which sampled entities come from different levels (e.g., work teams and employees). Moreover, collecting samples composed of higher-level units (e.g., teams, departments, organizations) is generally more demanding and difficult than collecting samples of individuals (e.g., employees). Taking these concerns into account, interested researchers frequently ask questions about what an adequate size is for a multilevel study sample. Two useful sources of information for making decisions about sample size recommendations when designing multilevel studies are: (a) The results of simulation studies in which the performance of multilevel methods under varied conditions (e.g., different numbers of units and individuals per unit) is examined, and (b) power analysis tools.

González-Romá and Hernández (2017) reviewed the aforementioned simulation studies published in recent decades. Based on their review, they suggested that when the focus is on testing cross-level direct (or Between-direct) effects and effect sizes are small-to-medium, unbiased estimates and a statistical power of at least .80 can be obtained with samples composed of 30 units and 20–40 individuals per group (Bell, Morgan, Schoeneberger, Kromrey, & Ferron, 2014). When the interest is in testing cross-level interactions, similar results for bias and power across different effect size conditions can be obtained with samples composed of 40 units and 18 individuals per group[5] (Mathieu, Aguinis, Culpepper, & Chen, 2012). As a last general guideline, "the studies reviewed do suggest that it is better to have more groups with fewer observations per group than the other way around" (González-Romá & Hernández, 2017, p. 193). Although helpful, these general recommendations and the more specific ones drawn from individual simulation studies must be interpreted with caution. They are based on the specific factors and conditions considered in simulations studies, which may not be generalizable to a given empirical study (Tonidandel, Williams, & LeBreton, 2015).

The information provided by simulation studies can be complemented with the information offered by power analysis tools specifically designed for multilevel studies. These tools allow researchers to ascertain what sample size combinations (number of units and individuals per unit) will reach acceptable statistical power values[6], taking into account the available information about different factors (e.g., ICC, reliability, effect sizes). These tools should be used when researchers are designing multilevel studies, before data are collected. There are tools focused on cross-level direct effects (Bosker, Snijders, & Guldemond, 2003; Browne, Lahi, & Parker, 2009) and cross-level interactions (Mathieu et al., 2012)[7]. Together, the results of simulation studies and the use of power analysis tools allow researchers to make informed decisions about sample size in multilevel designs, in addition to considering the practical constraints imposed by research work conducted with real work units (e.g., teams, departments, organizations).

---

[5]Here I only provide some basic guidelines for Conventional Multilevel Modeling methods. More detailed information about the performance of these methods in the studies reviewed by González-Romá and Hernández (2017) can be obtained via the following Supplemental Material link (https://www.annualreviews.org/doi/suppl/10.1146/annurev-orgpsych-041015-062407). González-Romá and Hernández (2017) also reviewed the performance of Structural Equation Modeling-Multilevel methods.

[6]See Cortina and Landis' (2009) for a discussion about factors to consider when establishing acceptable statistical power values.

[7]In these references, readers can find examples of statistical power calculations.

Conclusion

The multilevel approach offers a more comprehensive view of complex organizational phenomena. To reap the fruits of this approach, the relationships involved in multilevel theoretical and research models have to be interpreted adequately, the nature of higher-level constructs and the processes that account for their emergence must be explained, and the design of empirical studies has to be based on informed decisions. By discussing three issues pertaining to these three points, I hope this article contributes to the improvement of multilevel research in Work and Organizational Psychology.

References

Ashforth B. E. (1985). Climate formation: Issues and extensions. *The Academy of Management Review*, *10*, 837–847. https://doi.org/10.2307/258051.

Barsade S. G., & Knight A. P. (2015). Group affect. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*, 21–46. https://doi.org/10.1146/annurev-orgpsych-032414-11131

Bell B. A., Morgan G. B., Schoeneberger J. A., Kromrey J. D., & Ferron J. M. (2014). How low can you go? An investigation of the influence of sample size and model complexity on point and interval estimates in two-level linear models. *Methodology*, *1*, 86–92.

Bosker R. J., Snijders T. A. B., & Guldemond H. (2003). *PINT (Power IN Two-level designs): Estimating standard errors of regression coefficients in hierarchical linear models for power calculations: User's manual* (Version 2.1). Groningen, The Netherlands: University of Groningen.

Browne W. J., Lahi M. G., & Parker R. M. A. (2009). *A guide to sample size calculations for random effect models via simulation and the MLPowSim Software Package*. Bristol, UK: University of Bristol.

Burstein L., Linn R. L., & Capell F . J. (1978). Analyzing multilevel data in the presence of heterogeneous within-class regressions. *Journal of Educational Statistics*, *3*, 347–383. https://doi.org/10.3102/10769986003004347

Cortina J. M., & Landis R. S. (2009). When small effect sizes tell a big story, and when large effect sizes don't. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 287–308). New York, NY: Routledge.

de Leeuw J., & Kreft I. (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics*, *11*, 57–85. https://doi.org/10.3102/10769986011001057

Enders C. K., & Tofighi D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*, 121–138. https://doi.org/10.1037/1082-989X.12.2.121.

Goldstein H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*, *73*, 43–56. https://doi.org/10.1093/biomet/73.1.43

González-Romá V., & Hernández A. (2014). Climate uniformity: Its influence on team communication quality, task conflict, and team performance. *Journal of Applied Psychology*, *99*, 1042–1058. https://doi.org/10.1037/a0037868.

González-Romá V., & Hernández A. (2017). Multilevel modeling: research-based lessons for substantive researchers. *Annual Review of Organizational Psychology and Organizational Behavior*, 4, 183–210. https://doi.org/10.1146/annurev-orgpsych-041015-062407.

González-Romá V., & Peiró J. M. (2014). Climate and culture strength. In B. Schneider & K. Barbera (Eds.), *The Oxford handbook of organizational climate and culture* (pp. 496–531). New York, NY: Oxford University Press.

González-Romá V., Peiró J. M., & Tordera N. (2002). An examination of the antecedents and moderator influences of climate strength. *Journal of Applied Psychology*, 87, 465–473. https://doi.org/10.1037/0021-9010.87.3.465.

Hitt M. A., Beamish P. W., Jackson S. E., & Mathieu J. E. (2007). Building theoretical and empirical bridges across levels: Multilevel research in management. *Academy of Management Journal*, 50, 1385–1399. https://doi.org/10.5465/amj.2007.28166219.

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

James L. R., Demaree R. G., & Wolf G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, 69, 85–98. https://doi.org/10.1037/0021-9010.69.1.85.

Kelly J. R. & Barsade S. G. (2001). Mood and emotions in small groups and work teams. *Organizational Behavior and Human Decision Processes*, 86, 99–130.

Klein K. J., Dansereau F., & Hall R. J. (1994). Levels issues in theory development, data-collection, and analysis. *The Academy of Management Review*, 19, 195–229. https://doi.org/10.5465/amr.1994.9410210745

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

LoPilato A. C., & Vandenberg R. J. (2015). The not-so-direct cross-level direct effect. In C. E. Lance & R. J. Vandenberg (Eds.), *More statistical and methodological myths and urban legends* (pp. 292–310). New York, NY: Routledge.

Mathieu J. E., Aguinis H., Culpepper S. A., & Chen G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology*, 97, 951–966. https://doi.org/10.1037/a0028380.

Muthén B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, 54, 557–585.

Preacher K. J, Zyphur M. J., & Zhang Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, 15, 209–233. https://doi.org/10.1037/a0020141.

Rentsch J. R. (1990). Climate and culture: Interaction and qualitative differences in organizational meanings. *Journal of Applied Psychology*, 75, 668–681. https://doi.org/10.1037/0021-9010.75.6.668.

Rousseau D. M. (1985). Issues of level in organizational research: multi-level and cross-level perspectives. *Research in Organizational Behavior*, 7, 1–38.

Tonidandel S., Williams E. B., & LeBreton J. M. (2015). Size mattersjust not in the way that you think. In C. E. Lance & R. J. Vandenberg (Eds.), *More statistical and methodological myths and urban legends* (pp. 162–183). New York, NY: Routledge.