**Building an Organizational Science of Behavioral Consistency: Comment on Katz-Navon,
Kark, and Delegach (2019)**

Jonas W. B. Lang
Department of Personnel Management, Work and Organizational Psychology
Ghent University
Henri Dunantlaan 2
9000 Ghent, Belgium
Jonas.lang@ugent.be
and
Department of Management
University of Exeter

Zen Goh
Department of Management
Monash Business School
900 Dandenong Road
Caulfield East
Victoria 3145, Australia
Zen.goh@monash.edu

Executive Summary

This commentary links Katz-Navon et al.'s findings on behavioral inconsistency to two extant literatures on (a) climate strength in organizational units, and (b) on individual adaptability across situations, both touching on the notion of consistency. The authors also suggest that Katz-Navon et al.'s findings on the middle area of scales could be extended using psychometric methods like item-response tree modeling techniques that split responses to Likert-scales into sub processes.

Building an Organizational Science of Behavioral Consistency: Comment on Katz-Navon, Kark, and Delegach (2019)

The series of studies reported by Katz-Navon, Kark, and Delegach (2019) documented a curvilinear relationship between other-rated transformational leadership (TF) and safety behavior in followers. Unlike earlier studies on curvilinear predictor-criterion relationships in the organizational literature (e.g., Pierce & Aguinis, 2013), average levels of a desirable leader characteristic (TF) are associated with less desirable outcomes in their research. Katz-Navon et al. suggest that the middle area of a scale frequently represents inconsistent behavior in leaders and that consistent leaders, i.e., consistently high or low TF, may frequently be the key to success as an explanation. In this brief commentary, we build on Katz-Navon et al.'s (2019) findings on the middle area of the transformational leadership scale and how this is frequently related to inconsistency perceptions in two ways. First, we put their findings into a larger context by linking it to two extant literatures on consistency. Second, we suggest that future organizational research could gain additional insights through the use of recently described item-response modeling techniques that split Likert-scales into sub processes.

Consistency in the Social Environment vs. Individual Consistency Across Situations

Consistency is often implicitly included in thinking about organizations, for instance, in Daniel Katz and Robert Kahn's suggestion in their classic work *The Social Psychology of Organizations* that the essence of an organization is "patterned" human behavior (Katz & Kahn, 1978). A more explicit discussion of the notion of consistency has taken place in the literature on organizational climate. Researchers have argued that climate strength, which indicates the consistency of unit-level constructs like ethical or safety climate, is frequently equally or more important as the actual level of climate and form an important source of additional information (Bliese & Britt, 2001; Schneider et al., 2002). In this literature, climate strength is typically

assessed through either the within-unit standard deviation (*SD*) across team members or within-group agreement (r_{wg}) coefficients (Bliese, 2000; LeBreton & Senter, 2008) and used as an additional main effect or moderator. When a consistent climate emerges, the work environment is more stable because the unit-members reactions and behaviors to the situation is similar even when the situation on which unit-members agreed, for instance, an unfair organization or bad leadership, is itself not necessarily positive. However, group members may be better able to deal with a consistent situation, even if it is unpleasant. Recently, the climate literature has begun to more actively study how climates emerge in organizational units over time in an effort to better understand how patterned human behavior emerges in organization, or, conversely what hampers the formation of consistent climates (Kozlowski, Chao, Grand, Braun, & Kuljanin, 2013; Lang, Bliese, & Adler, 2019; Lang, Bliese, & de Voogt, 2018). The Katz-Navon et al. (2019) article is related to the literature on climate strength in the sense that leader characteristics can be seen as a characteristic of the environment much like consensus on the environment by unit-members and that arguably stable behavior should have advantages in the sense that employees can expect a stable and consistent environment.

The similarities between consistency in the Katz-Navon et al. (2019) paper and the climate literature notwithstanding, there are also important differences. Transformational leadership is conceptually and in its original description related to personality characteristics that are frequently seen as a distant antecedent of transformational leadership in addition to other elements (Cavazotte, Moreno, & Hickmann, 2012; Judge, Bono, Ilies, & Gerhardt, 2002). Interestingly, the abovementioned notion that consistency in behavior is positive as an environmental characteristic somewhat contrasts with ideas in research on individual-level consistency in personality, and also personality in work settings. Personality researchers have long been interested in variability in personality and have used various constructs like character

steadiness (Walton, 1936), personality strength (Dalal et al., 2015), personality flexibility (Paulhus & Martin, 1988), flux (Moskowitz & Zuroff, 2004) or variability (Baird, Le, & Lucas, 2006; Eid & Diener, 1999; Fiske & Rice, 1955; Fleeson, 2001; Lang, Lievens, De Fruyt, Zettler, & Tackett, 2019; Lievens et al., 2018) to capture this notion of consistency. More recently, several theoretical frameworks have emerged that incorporate variability (Carver & Scheier, 1998; Fleeson & Jayawickreme, 2015; Sosnowska, Kuppens, De Fruyt, & Hofmans, 2019). The most widely known of these frameworks is whole trait theory that describes personality as frequency distributions of states and essentially suggest that traits have not only a mean level but can also be defined by their variability (Fleeson, 2001; Fleeson & Jayawickreme, 2015; Lievens et al., 2018). The most common design used in this literature to measure variability are experience sampling designs or diary designs that ask participants to fill out one or more state personality reports (e.g., “Currently, I am sociable”) each day over multiple days (typically 5-15 days) and then estimate the *SD* in these reports for each person (Baird et al., 2006; Fleeson, 2001). The management literature has recently begun to study the degree to which variability in traits predicts behavior in organizational settings and job performance and has typically found that variability in state personality and item-response theory (IRT) variability is frequently associated with performance (Dalal et al., 2015; Lievens et al., 2018). A common explanation is that people with higher variability are more flexible or adaptable in their behavior, and the implicit takeaway is that flexibility in behavior can be positive for individuals in that one can adapt and adjust to situation demands sufficiently to complete the tasks.

The findings in the Katz-Navon et al. (2019) paper highlight the theoretical notion of consistency. We extended this discussion by linking it two extant literatures that also focus on the notion of consistency. While both literatures and the Katz-Navon et al. (2019) study discuss consistency, it becomes obvious upon closer observation that there is a contrast between the

notion that consistency is good for a social environment in both the climate literature and the findings of the Katz-Navon et al. (2019) study on the one hand, and the idea that inconsistency across situations captures a functional element in the personality and HRM literature on flexibility or adaptability on the other. How can those two opposing ideas be reconciled? One potential theoretical approach could be to think of inconsistency as a stressor when it is a characteristic of the environment and a resource at the individual level that allows organizational members to react flexibly and adaptively to their environment. Leadership behavior has an interesting double-function, as the leader is at the same time an organizational member and a characteristic of other organizational members' environment. Thus, consistency potentially captures an important element of tension or paradox in leadership as such. On the one hand, successful organizational members who get promoted may show flexibility and variability in their behavior. On the other hand, exactly this variability or inconsistency may not be what leadership positions require to create the type of stable environments that followers desire. Likewise, especially in modern environments, change may frequently be inevitable at some point requiring variability and flexibility to avoid failure. However, successful leaders may frequently especially be characterized by consistent and somewhat inflexible behavior that makes them predictable and successful with their followers. This consistency tension may be one puzzle in better understanding the typical high rate of leader failure.

The Middle Area of a Scale and Consistency: Psychometric Perspectives

Katz-Navon et al. link perceptions of inconsistency to the middle area of the TF scale. In contrast, extant literatures have typically operationalized a consistency/variability/strength trait dimension using *SDs*. However, one challenge in so doing is that individuals with extreme values on the scale—very high or very low—inevitably have a restricted amount of possible variability on the *SD*. An important question is thus how future research could gain additional insights into

the nature of the middle area of scales and inconsistency beyond the Katz-Navon et al. study that used an external measurement and the limitations of the *SD*/mean methodology.

We suggest that one potential way to address this measurement challenge is a recently developed psychometric perspective on Likert rating scales. This perspective was developed from the observation that there are two extreme ways to get an average score on a scale: In one, the person would endorse options around the middle of the scale all the time (2 = agree, 3 = undecided, and 4 = disagree). In the other, the person would mostly use the most extreme options (1 and 5) intermittently. Both persons could get the same score. To capture the difference between these response patterns, researchers can use a psychometric model—the trait variability tree model (TVTM; Lang, Lievens, et al., 2019; Lievens et al., 2018)—to measure IRT variability in responding. The TVTM is an IRT-based model that builds on a specific type of IRT models: IR tree models (Böckenholt, 2012; De Boeck & Partchev, 2012; LaHuis, Blackmore, Bryant-Lees, & Delgado, 2018). IR tree models are based on the idea that a response decision can have multiple underlying processes (also called pseudoitems). The TVTM builds on the observation that midpoint responding (mid-category yes=0 or no=1) and extreme responding (strongly agree/disagree = 0 or agree/disagree = 1 with the middle category coded as missing) are typically highly correlated with each other, but are both not necessarily correlated with the direction of responses to Likert items (strongly agree/agree = 1 or strongly disagree/disagree = 0 with the middle category coded as missing). It is thus possible to estimate both individual differences in people's content-related behavior (direction) and individual differences in a content-free IRT variability dimension (midpoint responding + extreme responding*[-1]), and to separate these behavioral tendencies from other response elements like measurement error, and careless responding.

The TVTM could be used in future organizational research on inconsistency/consistency

using Likert scales in a couple of ways. For instance, the approach presents an alternative for measuring inconsistency in future organizational research on leadership, safety, and in other areas so that researchers could study linear and curvilinear effects of direction responding and effects of IRT variability on organizational outcomes. Alternatively, the model could also allow some more fine-grained insights into how findings like that of the Katz-Navon et al. study emerge on the item-level.

Toward an Organizational Science of Behavioral Consistency

Katz-Navon et al. started a conversation about the meaning of the middle area of scales and the status of consistency in organizational research, with regards to safety behavior/performance. We believe that this conversation applies to different outcome variables (creativity, service, quality, etc.), and is important because the notion of consistency lies at the heart of the definition of organizations as patterned human behavior. We see several important steps for building a systematic organizational science of behavior consistency in future research. First, future research could systematically compare different theoretical ideas of consistency/inconsistency and explore their link to the middle area of scales. Second, it seems important to explore the implications of different operationalizations or measurement approaches for different theoretical conceptualizations of consistency and the relationships between these different approaches. Our discussion showed some disparities in how consistency is operationalized and measured in different literatures and we described how IRT methods could provide some novel insights in future research. One potentially important step could also be to systematically explore how consistency in behavior actually emerges in both individuals and groups in organizations over time to better understand the underlying mechanisms. Finally, and third, future research should likely tackle level of analysis issues by systematically studying consistency at different levels and top-down relationship between consistency and behavior in an extended multilevel framework.

References

- Baird, B. M., Le, K., & Lucas, R. E. (2006). On the nature of intraindividual personality variability: Reliability, validity, and associations with well-being. *Journal of Personality and Social Psychology*, 90, 512–527. <https://doi.org/10.1037/0022-3514.90.3.512>
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.
- Bliese, P. D., & Britt, T. W. (2001). Social support, group consensus and stressor-strain relationships: social context matters. *Journal of Organizational Behavior*, 22(4), 425–436. <https://doi.org/10.1002/job.95>
- Böckenholt, U. (2012). Modeling multiple response processes in judgment and choice. *Psychological Methods*, 17, 665–678. <https://doi.org/10.1037/a0028111>
- Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York, NY: Cambridge University Press.
- Cavazotte, F., Moreno, V., & Hickmann, M. (2012). Effects of leader intelligence, personality and emotional intelligence on transformational leadership and managerial performance. *The Leadership Quarterly*, 23(3), 443–455. <https://doi.org/10.1016/j.leaqua.2011.10.003>
- Dalal, R. S., Meyer, R. D., Bradshaw, R. P., Green, J. P., Kelly, E. D., & Zhu, M. (2015). Personality Strength and Situational Influences on Behavior. *Journal of Management*, 41(1), 261–287. <https://doi.org/10.1177/0149206314557524>
- De Boeck, P., & Partchev, I. (2012). IRTrees: Tree-based item response models of the GLMM family. *Journal of Statistical Software*, 48(1), 1–18. <https://doi.org/10.18637/jss.v048.c01>
- Eid, M., & Diener, E. (1999). Intraindividual variability in affect: Reliability, validity, and

personality correlates. *Journal of Personality and Social Psychology*, 76, 662–676.

<https://doi.org/10.1037//0022-3514.76.4.662>

Fiske, D. W., & Rice, L. (1955). Intra-individual response variability. *Psychological Bulletin*, 52, 217–250. <https://doi.org/10.1037/h0045276>

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, 80, 1011–1027. <https://doi.org/10.1037//0022-3514.80.6.1011>

Fleeson, W., & Jayawickreme, E. (2015). Whole trait theory. *Journal of Research in Personality*, 56, 82–92. <https://doi.org/10.1016/j.jrp.2014.10.009>

Greene, W. H. (2012). *Econometric analysis* (7th ed.). Essex, UK: Pearson.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87(4), 765–780. <https://doi.org/10.1037/0021-9010.87.4.765>

Katz-Navon, T., Kark, R., & Delegach, M. (2019). Trapped in the middle: Challenging the linear approach to the relationship between leadership and safety. *Academy of Management Discoveries*.

Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations*. New York, NY, US: Wiley.

Kozlowski, S. W. J., Chao, G. T., Grand, J. A., Braun, M. T., & Kuljanin, G. (2013). Advancing multilevel research design: Capturing the dynamics of emergence. *Organizational Research Methods*, 16, 581–615. <https://doi.org/10.1177/1094428113493119>

LaHuis, D. M., Blackmore, C. E., Bryant-Lees, K. B., & Delgado, K. (2018). Applying Item Response Trees to Personality Data in the Selection Context. *Organizational Research Methods*, 109442811878031. <https://doi.org/10.1177/1094428118780310>

- Lang, J. W. B., Bliese, P. D., & Adler, A. B. (2019). Opening the black box: A multilevel framework for studying group processes. *Advances in Methods and Practices in Psychological Science*. <https://doi.org/10.1177/2515245918823722>
- Lang, J. W. B., Bliese, P. D., & de Voogt, A. (2018). Modeling consensus emergence in groups using longitudinal multilevel models. *Personnel Psychology*, 71, 255–281. <https://doi.org/10.1111/peps.12260>
- Lang, J. W. B., Lievens, F., De Fruyt, F., Zettler, I., & Tackett, J. L. (2019). Assessing meaningful within-person variability in Likert-scale rated personality descriptions: An IRT tree approach. *Psychological Assessment*, 31, 474–487. <https://doi.org/10.1037/pas0000600>
- LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 Questions and Interrater Agreement. *Organizational Research Methods*, 11, 815–852. <https://doi.org/10.1177/1094428106296642>
- Lievens, F., Lang, J. W. B., De Fruyt, F., Corstjens, J., Van De Vijver, M., & Bledow, R. (2018). The predictive power of people's intra-individual variability across situations: Implementing whole trait theory in assessment. *Journal of Applied Psychology*, 103, 753–771. <https://doi.org/10.1037/apl0000280>
- McFadden, D. (2001). Economic choices. *American Economic Review*, 91, 351–378. <https://doi.org/10.1257/aer.91.3.351>
- Moskowitz, D. S., & Zuroff, D. C. (2004). Flux, pulse, and spin: Dynamic additions to the personality lexicon. *Journal of Personality and Social Psychology*, 86, 880–893. <https://doi.org/10.1037/0022-3514.86.6.880>
- Paulhus, D. L., & Martin, C. L. (1988). Functional flexibility: A new conception of interpersonal flexibility. *Journal of Personality and Social Psychology*, 55(1), 88–101. <https://doi.org/10.1037/0022-3514.55.1.88>

Pierce, J. R., & Aguinis, H. (2013). The Too-Much-of-a-Good-Thing Effect in Management.

Journal of Management, 39(2), 313–338. <https://doi.org/10.1177/0149206311410060>

Schneider, B., Salvaggio, A. N., & Subirats, M. (2002). Climate strength: a new direction for climate research. *Journal of Applied Psychology*, 87(2), 220–229.

<https://doi.org/10.1037/0021-9010.87.2.220>

Sosnowska, J., Kuppens, P., De Fruyt, F., & Hofmans, J. (2019). A dynamic systems approach to personality: The Personality Dynamics (PersDyn) model. *Personality and Individual Differences*, 144, 11–18. <https://doi.org/10.1016/j.paid.2019.02.013>

Walton, R. D. (1936). Relations between amplitude of oscillations in short period efficiency and steadiness of character. *British Journal of Psychology*, XXVII.