Back to Basics: The Importance of Conceptual Clarification in Psychological Science

Laura F. Bringmann[1,2] , Timon Elmer[1], and Markus I. Eronen[3]

[1]Department of Psychometrics and Statistics, Faculty of Social and Behavioural Sciences, University of Groningen; [2]Interdisciplinary Center Psychopathology and Emotion regulation (ICPE), Department of Psychiatry, University Medical Center Groningen, University of Groningen; and [3]Department of Theoretical Philosophy, University of Groningen

Abstract

Although the lack of conceptual clarity has been observed to be a widespread and fundamental problem in psychology, conceptual clarification plays a mostly marginal role in psychological research. In this article, we argue that better conceptualization of psychological phenomena is needed to move psychology forward as a science. We first show how conceptual unclarity seeps through all aspects of psychological research, from everyday concepts to statistical measures. We then turn to recommendations on how to improve conceptual clarity in psychology, emphasizing the importance of seeing research as an iterative process in which it is necessary to revisit the phenomena that are the foundations of theories and models, as well as how they are conceptualized and measured.

Keywords

concept, phenomena, measurement, qualitative and quantitative methods, philosophy of science

> Premature advanced science stifles creativity, closes the eyes of the field to important new phenomena, is prone to generate long lines of research that ultimately have little to do with the basic target of the field . . . and generally pulls people prematurely away from the real world, where it all starts. (Rozin, 2001, p. 5)

No psychological researcher would deny that "good models are hard to build on the basis of bad data" (Freedman, 1985, p. 345). However, what is often overlooked is that in order to get good data, psychological concepts must be clearly defined and measured in a valid way (Flake & Fried, 2020). Concepts are often referred to as ideas or terms that are the "building blocks" of thoughts and theories (Gerring, 1999; Podsakoff et al., 2016), and thus play a key role in all aspects of psychological research. Although the lack of conceptual clarity has been observed to be a widespread and fundamental problem in psychology (Antonakis, 2017; Eronen & Bringmann, 2021; Flake, 2021; MacKenzie, 2003; Podsakoff et al., 2016; Scheel et al., 2021), conceptual clarification plays a mostly marginal role in psychological research (see, e.g., Aguinis & Vandenberg, 2014).

It is important not to conflate conceptual clarification with construct validity, which primarily concerns whether a test measures the construct it is intended to measure (Borsboom et al., 2004). In contrast, conceptual clarification is about characterizing the concept itself, which can be independent of measurement (see also Cartwright, 2009). For example, phenomena such as fear of spiders can be observed, described, and conceptualized without determining how to measure the fear of spiders. In general, as we emphasize, research usually proceeds in iterative cycles in which all aspects (observation of phenomena, conceptualization, measurement, and theory) are intertwined.

Corresponding Author:
Laura F. Bringmann, Department of Psychometrics and Statistics, Faculty of Social and Behavioural Sciences, University of Groningen
Email: l.f.bringmann@rug.nl

In this article, we argue that better conceptualization of psychological phenomena is needed to move psychology forward as a science. Concepts are the necessary elements of high-quality theories, methods, and data (Podsakoff et al., 2016; Scheel et al., 2021), and without them, psychological science will lack a solid foundation, which will hamper scientific progress (Gerring, 1999; Wilshire et al., 2021). In what follows, we first highlight the importance of conceptual clarity and then elaborate on different ways in which to improve conceptual clarity in psychology. To exemplify the consequences of the absence of conceptual clarity for psychological science and ways forward to improve the research of our discipline, we draw on concepts from our own research, namely, friendship, psychopathological symptoms, and centrality. However, our arguments are not restricted to these cases, but applicable more broadly to all fields of psychology.

The Importance of Conceptual Clarity

In psychology and social sciences in general, many of the concepts used in research, such as "friendship," "emotions," or "concentration," come from everyday life and therefore may not seem to be in need of clarification. However, when one tries to define a concept such as "friendship," it becomes clear that people understand it in widely different ways, and studies show that whom people call a friend is often even inconsistent, because of differences in how "friendship" is interpreted in different contexts (Fischer, 1982). Although the lack of clarity may not matter much when concepts are used in everyday conversations, conceptual clarity is essential when the aim is to apply these concepts in a systematic way—in other words, to do research. The need for conceptual clarity is broadly acknowledged, but nevertheless, concepts are usually not precisely and clearly defined in empirical research, even when they are the main topic of study (for an example of such lack of clarity regarding "friendship," see, e.g., Feiler & Kleinbaum, 2015).

Even specialized concepts whose meaning has far-reaching societal and medical consequences often lack conceptual clarity. Consider the case of psychological symptoms. Psychological symptoms are the building blocks of most frameworks for studying or classifying mental disorders—ranging from the *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association, 2013) to the latest transdiagnostic frameworks, such as the psychological network approach (Borsboom & Cramer, 2013) or dimensional models (e.g., the Hierarchical Taxonomy of Psychopathology, or HiTOP; Kotov et al., 2017)—and thus are essential for diagnosing and treating people with mental disorders.

At first glance, what a symptom is and how a symptom should be measured may seem evident. However, even with regard to well-established symptoms, such as impaired concentration (as a symptom of, e.g., depression), many conceptual questions arise.

For instance, as Wilshire et al. (2021) pointed out, studies of patients' narratives suggest that what is now called impaired concentration may turn out to be several distinct phenomena. Whereas some patients describe the experience of impaired concentration as "blanking," others describe it as the feeling of concentration being interrupted by intrusive thoughts, and yet others describe it as an experience of the mind drifting off topic. Wilshire et al. further emphasized that this example is not an exception, as similar heterogeneity and conceptual ambiguity can be found in many other canonical symptoms, such as anhedonia and fatigue. Such ambiguity casts serious doubts on the validity of symptom measurements: If the interpretation of impaired concentration can vary widely across researchers and participants, the meaning of an item purported to measure this construct is ambiguous, and it is unclear what such an item actually measures.

These conceptual issues also undermine the foundations of theories and statistical models of mental disorders, or as Wilshire et al. (2021) put it: "There is little point in continuing to develop and refine statistical techniques or classification schemes until we have a better grasp of these key concepts" (p. 336; see also Jacobucci & Grimm, 2020). If impaired concentration consists of several distinct phenomena, but is represented by just one variable in theoretical or statistical models of depression, then those models mischaracterize the structure and dynamics of depression in a fundamental way. The interrelations among the different phenomena that are lumped together as impaired concentration, as well as the specific relationships each of these phenomena has with other symptoms of depression, are not represented in the models. Thus, the theoretical and statistical models based on this ambiguous conceptualization do not accurately capture the basic structure of the system, which hampers understanding and treatment of the disorder.

These issues are the very basis of Freedman's (1985) concern that "bad data" make it hard to build useful and meaningful theories or statistical models. Moreover, these problems are deeply conceptual, and not just statistical: Statistics on its own can never reveal whether concepts (and therefore the models built upon them) are well defined and in line with the research question (see also Borsboom et al., 2004). For that, the interpretation and conceptual work of the researcher are needed, and one cannot solely rely on numeric results, as can be illustrated by the use of factor analysis. When

items load on the same factor, it is appealing to interpret this as showing that they measure the same phenomenon. However, as Maul (2017) showed, even completely made-up nonsense data can result in a neat one-factor (or two-factor) solution with high loadings and correlations with other variables. Therefore, factor analysis cannot indicate whether high factor loadings are due to items actually measuring the same phenomenon or due to (unintended) semantic or conceptual overlap between the items. As Borsboom et al. (2004) pointed out, "no amount of empirical data can fill a theoretical gap" (p. 1068).

Moreover, conceptual unclarity easily seeps through statistical analyses as well, as statistical measures are often intertwined with conceptualizations (see also Gelman & Hennig, 2017). For example, consider centrality measures, which are often used in social-networks research, and increasingly also in psychological research, where they are applied to psychological network models (i.e., models representing partial correlations between affective states or symptoms; Borsboom & Cramer, 2013). At first glance, centrality seems to be an intuitive concept that can be investigated using centrality measures indicating the most important node (e.g., person or symptom) in a network. Mathematically speaking, centrality measures are derived from a node's position in a network; for example, a node's distance from other nodes in a network can be calculated to obtain a measure of its closeness centrality (Freeman, 1979). Such distance-based centrality measures have been widely applied in psychological networks to estimate a symptom's influence on the system (between 2008 and 2018, at least 145 articles reported results using these measures; Robinaugh et al., 2020).

Given a precise mathematical definition for a node's centrality, one might be tempted to interpret this value as a precise measure of the node's importance. Indeed, researchers (including the first author of this article; Bringmann et al., 2016) did not initially realize that using distance-based centrality measures in correlation and partial correlation networks is conceptually flawed. In networks that, for example, represent the connections between train stations, applying such a distance-based centrality measure makes conceptual sense. However, this reasoning does not work for a correlation or partial correlation network. In this case, the connections represent how strongly the nodes are associated, not how far or close they are from each other, so adding up the connection values does not result in a measure of distance (Bringmann et al., 2019). Although the mathematical calculation of the centrality measure is straightforward, this measure comes with an implicit conceptualization of what centrality is, and this conceptualization simply does not match with the models and data that psychological networks are based on.

Similarly, conceptual issues cannot be avoided when statistical techniques such as predictive algorithms (e.g., machine learning) are used. It may seem that predictive algorithms are free from assumptions, as they simply allow one to make predictions about certain outcomes. However, their application always already involves (often implicitly) assuming a certain conceptualization of the predictors and outcome (see also Jacobucci & Grimm, 2020). For instance, consider studies aimed at determining the best predictors of successful treatment outcomes for patients with depression. What it means for treatment of depression to be successful has been hotly debated, and successful treatment has been conceptualized in many alternative ways (Fava et al., 2007). If treatment success in a predictive study is defined in a way that does not match at all with the experiences of patients or practitioners (see Slofstra et al., 2019), all the predictions will be rendered meaningless and unhelpful for actual practice, no matter how good the predictions are.

Current Directions for Improving Conceptual Clarity

In the previous section, we showed that conceptual clarity (or unclarity) is not just something that involves concepts and their definitions, but rather influences all aspects of research, from measurement to the statistical methods used to answer a research question. Given its fundamental importance, we now turn to suggestions on how to improve the conceptual foundations of psychological research.

Such recommendations should be straightforward and easy to implement for a broad range of researchers. Although we find existing discussions of conceptualization insightful, and largely build on them in this article, they often have focused too much on providing long lists of rather abstract criteria for conceptual clarity, which can set the bar very high for starting conceptual-clarification research or clarifying a concept in an empirical article (e.g., Gerring, 1999; Podsakoff et al., 2016). Instead, we believe that the focus should be on step-by-step epistemic iteration, so that not all criteria have to be met at once.

How such epistemic iteration operates can be illustrated with an example from the history of science. As extensively discussed by Chang (2004), the measurement of temperature started from very rough conceptualizations based on the subjective sensations of warm and cold (e.g., feeling how hot the water is with one's hand). This formed the basis for simple measurement

Table 1. Recommendations for Improving Conceptual Clarity

Recommendations for research articles
    1. In research, go through the whole iterative cycle, and thus go back to basics (e.g., conceptualization)
    2. Explicitly discuss conceptual ambiguities and different ways of defining the key concepts
    3. Link concepts to the measurement methods used and justify how the measurements capture the concepts
Recommendations for psychology in general
    4. Conduct research solely focusing on specific concepts (i.e., conceptual studies)
    5. Dedicate funds for research on conceptual clarification
    6. Incorporate conceptual competence in psychological education
    7. Teach qualitative research skills in psychological education
    8. Place value on and enhance the visibility of commentary that addresses conceptual flaws in published studies

instruments (e.g., first thermoscopes, later thermometers), and on the basis of those measurements, the concepts (e.g., temperature, heat) were refined and redefined. This led to better instruments, which eventually made possible the development of new theories and statistical procedures (e.g., statistical mechanics), which in turn helped to make the concepts more well defined by embedding them in a broader theoretical framework (Bringmann & Eronen, 2016; Chang 2004). Thus, science advances in iterative cycles in which description of phenomena, concept definition and refinement, measurement, statistics and theorizing are intertwined (see also Kendler, 2009, and the cycle proposed by Valsiner, 2014).[1]

We suggest that psychological research too often gets stuck in one part of this cycle, and that researchers should more frequently go through the whole iterative cycle and in this way keep going back to "the basics" (Recommendation 1 in Table 1, which lists all the recommendations we discuss in this section). Thus, when using theoretical or statistical models, one should revisit the phenomena that are the foundations of the models (see also Bogen & Woodward, 1988; Haig, 2013), as well as how they are conceptualized and measured. A concrete way of implementing this recommendation would be to explicitly discuss these issues in empirical research articles (see also Flake & Fried, 2020). In the Method section (or elsewhere), it is not enough to describe statistical methods and give an operationalization of the key concepts; in addition, conceptual ambiguities and different ways of defining the key concepts should be explicitly addressed (Recommendation 2). In addition, the definitions should be clearly linked to the measurement procedures used, and researchers should provide justification for why these measurement procedures are good ways of capturing the concepts that have been defined (Recommendation 3; see also Cartwright, 2009; Flake & Fried, 2020).

To make this more specific, we return to the friendship example. One article reporting a study of friendship defined friends as "those individuals who you feel close to, who you interact with frequently, those who you would seek out to do some type of social activity" (Brewer & Webster, 2000, p. 363). Even though it is laudable that the authors gave an explicit definition, which is also not always done, the definition was vague: What does it mean to feel close to an individual? When are interactions frequent? Is one interaction per day required, or is one per week or month sufficient? And what counts as a social activity? Is chatting after a work meeting a social activity? These questions are essential for studying friendship. For example, every participant might interpret the quoted definition differently, which would make measurements incomparable between persons.[2] Although these conceptual unclarities have not remained unnoticed in the literature on friendship (see, e.g., Fischer, 1982; Kitts & Leal, 2021), vague definitions and measures are still often used, and thus conceptual issues in applying measures of friendship remain unresolved.

Going beyond recommendations for empirical research procedures and the clarification of concepts in manuscripts, we further recommend shifts in how psychological science is approached on a more general level. Specifically, it should be standard that some articles and studies focus on specific concepts: how they are defined, how they are measured, and how the definitions and measurements could be improved to represent real-world psychological phenomena of interest (Recommendation 4). Thus, instead of immediately studying relationships between two concepts, it would be essential to conduct research projects focusing solely on one concept at a time (see Clack & Ward, 2020, for an example of a project focusing specifically on anhedonia). In such a project, researchers could compare and improve definitions of the concept and how it is measured, and conduct qualitative interviews in order to evaluate how participants interpret the concept, whether participants generally have similar semantic boundaries for it (i.e., what phenomena are included and excluded in the concept), or the extent to which the concept differs from closely related concepts (see

also Kitts & Leal, 2021; Podsakoff et al., 2016; Tafreshi et al., 2016).

Such conceptualization studies could furthermore be strengthened by building a multiverse analysis into the study design. A multiverse analysis examines systematically how research results are influenced by the many choices that researchers need to make (Steegen et al., 2016). So far, multiverse analyses have been applied mainly to choices in data preparation, modeling, and statistical analysis, but the same approach can also be applied to conceptual choices. In the case of friendship, one could analyze, for example, if and how alternative ways of defining "feeling close" change the results (i.e., the number of friendships participants report having). (See also Huth et al., 2022, who applied a similar approach to a study on the relationship between large cities and depression and showed that different definitions of "city" lead to different results.)

Furthermore, dedicated funding for research on conceptual clarity would help in creating opportunities to do this kind of research (Recommendation 5), as has been done for replication studies (e.g., funding by the National Science Foundation in the United States and NWO in The Netherlands; see Cook, 2016, and NWO, 2022). Such funding would be important because conceptual research does not necessarily lead directly to new empirical findings and is therefore difficult to "sell" to journals and funding agencies, but is nevertheless the very basis for fruitful empirical research (Aguinis & Vandenberg, 2014; Freedman, 1985).

In addition, it would benefit the profession if psychologists were trained in "conceptual competence" (Aftab & Waterman, 2021), which includes philosophy of science, history and philosophy of one's field, and conceptual tools such as thinking and arguing in an organized and structured manner, using, for example, logic and argumentation theory (Recommendation 6). Such training can, among other things, help in choosing fruitful research questions and identifying implicit assumptions in theories, measurement, and data. Reflecting on key assumptions is also crucial for empirical research, as it allows one to see when statistical analyses are and are not applicable (see the previous section and the example on centrality).

Psychology students should also receive more training in analyzing qualitative information, which is of key importance for improving conceptual clarity (Recommendation 7; see also Tafreshi et al., 2016). Qualitative information from structured interviews, for example, can help in studying how participants understand a certain concept. Also, the open text boxes that are often included in questionnaires (e.g., in experience-sampling studies; Myin-Germeys & Kuppens, 2022) provide opportunities for participants to describe, for instance, what they have experienced and why they feel a certain way. Such responses potentially give a wealth of information, which can help researchers understand how participants interpret items and whether the researchers' own interpretation matches participants' (see also Truijens et al., 2022, Figs. 2 and 3, for an illustration of a patient's added annotations to a quantitative questionnaire). However, as handbooks and guidelines focusing on quantitative research usually do not even discuss qualitative data (see, e.g., Myin-Germeys & Kuppens, 2022), psychological researchers often do not have a clear idea of how such data could be analyzed and put to use.

Conceptual research can also help to guide the limited financial and labor resources in psychology to fruitful avenues. For example, replication research has received much attention in psychology in recent years because of the replication crisis, but considering the problems with ill-defined concepts discussed in the previous section, we argue that conceptual clarity is of key importance in deciding which studies to replicate. If concepts central to a study are very unclear and badly defined, or even fundamentally flawed, it may be advisable not to waste valuable resources on replicating that study, because it is already flawed on a conceptual basis (see also Flake, 2021). An example of such a concept is "psi," a central concept in parapsychology, which has so far not been connected to reasonable observable phenomena (see, e.g., Reber & Alcock, 2020). More effective than replication in such cases would be placing value on and enhancing the visibility of commentary that addresses the conceptual shortcomings (Recommendation 8; for an example of such commentary, see Draheim et al., 2019).

Conclusion

Even though conceptualization is intertwined with the rest of scientific research, it is a crucial part of research that should get more attention, in order to give psychology the solid basis it needs to develop as a science.

Recommended Reading

Aftab, A., & Waterman, G. S. (2021). (See References). Discusses the importance of conceptual thinking for clinical psychology and introduces the notion of conceptual competence.

Eronen, M. I., & Bringmann, L. F. (2021). (See References). Discusses the reasons why there are so few good theories in psychology, emphasizing the importance of clear conceptualization.

Rozin, P. (2001). (See References). Argues that psychological research is too focused on advanced methods at the expense of observing and establishing basic phenomena.

Wilshire, C. E., Ward, T., & Clack, S. (2021). (See References). Describes the conceptual problems underlying descriptions of psychological symptoms in a clear and thorough way.

Transparency

ORCID iDs

Laura Bringmann (iD) https://orcid.org/0000-0002-8091-9935
Markus Eronen (iD) https://orcid.org/0000-0003-2028-3338

Acknowledgments

Notes

1. As all the aspects of epistemic iteration are thoroughly intertwined, the idea of an iterative "cycle" should be understood metaphorically, not literally.
2. The conceptual ambiguity of friendship measurements is possibly responsible for the observation that when adolescents report who their friends are, only about 50% of those reports are reciprocated (i.e., only about half the time does the person named also consider the person making the report a friend; Vörös et al., 2019).

References

Aftab, A., & Waterman, G. S. (2021). Conceptual competence in psychiatry: Recommendations for education and training. *Academic Psychiatry*, *45*(2), 203–209. https://doi.org/10.1007/s40596-020-01183-3

Aguinis, H., & Vandenberg, R. J. (2014). An ounce of prevention is worth a pound of cure: Improving research quality before data collection. *Annual Review of Organization Psychology and Organizational Behavior*, *1*, 569–595. https://doi.org/10.1146/annurev-orgpsych-031413-091231

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *Leadership Quarterly*, *28*(1), 5–21. https://doi.org/10.1016/j.leaqua.2017.01.006

Bogen, J., & Woodward, J. (1988). Saving the phenomena. *The Philosophical Review*, *97*(3), 303–352. https://doi.org/10.2307/2185445

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91–121. https://doi.org/doi:10.1146/annurev-clinpsy-050212-185608

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*(4), 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

Brewer, D. D., & Webster, C. M. (2000). Forgetting of friends and its effects on measuring friendship networks. *Social Networks*, *21*(4), 361–373. https://doi.org/10.1016/S0378-8733(99)00018-0

Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., Wigman, J. T. W., & Snippe, E. (2019). What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology*, *128*(8), 892–903. https://doi.org/10.1037/abn0000446

Bringmann, L. F., & Eronen, M. I. (2016). Heating up the measurement debate: What psychologists can learn from the history of physics. *Theory and Psychology*, *26*(1), 27–43. https://doi.org/10.1177/0959354315617253

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., Tuerlinckx, F., & Kuppens, P. (2016). Assessing temporal emotion dynamics using networks. *Assessment*, *23*(4), 425–435. https://doi.org/10.1177/1073191116645909

Cartwright, N. D. (2009). What is this thing called efficacy? In C. Mantzavinos (Ed.), *Philosophy of the social sciences: Philosophical theory and scientific practice* (pp. 185–206). Cambridge University Press.

Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford University Press.

Clack, S., & Ward, T. (2020). Modeling the symptoms of psychopathology: A pluralistic approach. *New Ideas in Psychology*, *59*, Article 100799. https://doi.org/10.1016/j.newideapsych.2020.100799

Cook, F. L. (2016, September 20). *Dear Colleague Letter: Robust and reliable research in the social, behavioral, and economic sciences* (NSF16-137). National Science Foundation. https://www.nsf.gov/pubs/2016/nsf16137/nsf16137.pdf

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508–535. https://doi.org/10.1037/bul0000192

Eronen, M. I., & Bringmann, L. F. (2021). The theory crisis in psychology: How to move forward. *Perspectives on Psychological Science*, *16*(4), 779–788. https://doi.org/10.1177/1745691620970586

Fava, G. A., Ruini, C., & Belaise, C. (2007). The concept of recovery in major depression. *Psychological Medicine*, *37*(3), 307–317. https://doi.org/10.1017/S0033291706008981

Feiler, D. C., & Kleinbaum, A. M. (2015). Popularity, similarity, and the network extraversion bias. *Psychological Science*, *26*(5), 593–603. https://doi.org/10.1177/0956797615569580

Fischer, C. S. (1982). What do we mean by 'friend'? An inductive study. *Social Networks*, *3*(4), 287–306. https://doi.org/10.1016/0378-8733(82)90004-1

Flake, J. K. (2021). Strengthening the foundation of educational psychology by integrating construct validation into open science reform. *Educational Psychologist*, *56*(2), 132–141. https://doi.org/10.1080/00461520.2021.1898962

Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, *3*(4), 456–465. https://doi.org/10.1177/2515245920952393

Freedman, D. A. (1985). Statistics and the scientific method. In W. M. Mason & S. E. Fienberg (Eds.), *Cohort analysis in social research* (pp. 343–366). Springer. https://doi.org/10.1007/978-1-4613-8536-3_11

Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks*, *1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Gelman, A., & Hennig, C. (2017). Beyond subjective and objective in statistics. *Journal of the Royal Statistical Society A: Statistics in Society*, *180*(4), 967–1033. https://doi.org/10.1111/rssa.12276

Gerring, J. (1999). What makes a concept good? A criterial framework for understanding concept formation in the social sciences. *Polity*, *31*(3), 357–393. https://doi.org/10.2307/3235246

Haig, B. D. (2013). Detecting psychological phenomena: Taking bottom-up research seriously. *The American Journal of Psychology*, *126*(2), 135–153. https://doi.org/10.5406/amerjpsyc.126.2.0135

Huth, K. B. S., Finnemann, A., van den Ende, M. W. J., & Sloot, P. M. A. (2022). No robust relation between larger cities and depression. *Proceedings of the National Academy of Sciences, USA*, *119*(2), Article e2118943118. https://doi.org/10.1073/pnas.2118943118

Jacobucci, R., & Grimm, K. J. (2020). Machine learning and psychological research: The unexplored effect of measurement. *Perspectives on Psychological Science*, *15*(3), 809–816. https://doi.org/10.1177/1745691620902467

Kendler, K. S. (2009). An historical framework for psychiatric nosology. *Psychological Medicine*, *39*(12), 1935–1941. https://doi.org/10.1017/S0033291709005753

Kitts, J. A., & Leal, D. F. (2021). What is(n't) a friend? Dimensions of the friendship concept among adolescents. *Social Networks*, *66*, 161–170. https://doi.org/10.1016/j.socnet.2021.01.004

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., . . . Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*(4), 454–477. https://doi.org/10.1037/abn0000258

MacKenzie, S. B. (2003). The dangers of poor construct conceptualization. *Journal of the Academy of Marketing Science*, *31*(3), 323–326. https://doi.org/10.1177/0092070303031003011

Maul, A. (2017). Rethinking traditional methods of survey validation. *Measurement*, *15*(2), 51–69. https://doi.org/10.1080/15366367.2017.1348108

Myin-Germeys, I., & Kuppens, P. (Eds.). (2022). *The open handbook of experience sampling methodology: A step-by-step guide to designing, conducting, and analyzing ESM studies* (2nd ed.). Center for Research on Experience Sampling and Ambulatory Methods Leuven.

NWO. (2022). *Replication studies.* https://www.nwo.nl/en/researchprogrammes/replication-studies

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2016). Recommendations for creating better concept definitions in the organizational, behavioral, and social sciences. *Organizational Research Methods*, *19*(2), 159–203. https://doi.org/10.1177/1094428115624965

Reber, A. S., & Alcock, J. E. (2020). Searching for the impossible: Parapsychology's elusive quest. *American Psychologist*, *75*(3), 391–399. https://doi.org/10.1037/amp0000486

Robinaugh, D. J., Hoekstra, R. H. A., Toner, E. R., & Borsboom, D. (2020). The network approach to psychopathology: A review of the literature 2008–2018 and an agenda for future research. *Psychological Medicine*, *50*(3), 353–366. https://doi.org/10.1017/S0033291719003404

Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review*, *5*(1), 2–14. https://doi.org/10.1207/S15327957PSPR0501_1

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why hypothesis testers should spend less time testing hypotheses. *Perspectives on Psychological Science*, *16*(4), 744–755. https://doi.org/10.1177/1745691620966795

Slofstra, C., Booij, S. H., Hoenders, H. J. R., & Castelein, S. (2019). Redefining therapeutic outcomes of depression treatment. *Journal for Person-Oriented Research*, *5*(2), 115–122. https://doi.org/10.17505/jpor.2019.10

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Tafreshi, D., Slaney, K. L., & Neufeld, S. D. (2016). Quantification in psychology: Critical analysis of an unreflective practice. *Journal of Theoretical and Philosophical Psychology*, *36*(4), 233–249. https://doi.org/10.1037/teo0000048

Truijens, F. L., Desmet, M., De Coster, E., Uyttenhove, H., Deeren, B., & Meganck, R. (2022). When quantitative measures become a qualitative storybook: A phenomenological case analysis of validity and performativity of questionnaire administration in psychotherapy research. *Qualitative Research in Psychology*, *19*(1), 244–287. https://doi.org/10.1080/14780887.2019.1579287

Valsiner, J. (2014). Needed for cultural psychology: Methodology in a new key. *Culture & Psychology*, *20*(1), 3–30. https://doi.org/10.1177/1354067X13515941

Vörös, A., Block, P., & Boda, Z. (2019). Limits to inferring status from friendship relations. *Social Networks*, *59*, 77–97. https://doi.org/10.1016/j.socnet.2019.05.007

Wilshire, C. E., Ward, T., & Clack, S. (2021). Symptom descriptions in psychopathology: How well are they working for us? *Clinical Psychological Science*, *9*(3), 323–339. https://doi.org/10.1177/2167702620969215