

# Double, Double, Measurement Trouble, Sorry to Burst your Construct Bubble

**Talea Cornelius**  
Columbia University Irving  
Medical Center, USA

**Paulina Schenk**  
University College London,  
UK

**Diane Dixon**  
Edinburgh Napier  
University, UK

**Geert Crombez**  
Universitat Gent, Belgium

**Marie Johnston**  
University of Aberdeen, UK

theoretical constructs that underlie behaviour change.

## Clear, Collaborative, and Cumulative: Goals for Behavioural Science

A collective focus on mechanisms of action (MoAs) that underlie behaviour change in the field of behavioural intervention research has great potential to not only improve the efficacy and scalability of our interventions, but also to tailor these interventions across populations and inform theory development (Sumner et al., 2018a, 2018b). Yet, we are hampered by our own history: a grab-bag of theories and associated constructs that are poorly defined, leaving behavioural scientists with the difficult task of identifying the shared and unique components of this work and struggling to measure theoretical constructs in consistent and

This article summarizes a roundtable discussion held at the 37th Annual Conference of the European Health Psychology Society in Bremen, Germany (September 2023), titled 'Double, Double, Measurement Trouble, Sorry to Burst your Construct Bubble.' The focus of this discussion was on valid measurement of

rigorous ways (Peters & Crutzen, 2022). Critically, poorly specified constructs and the resultant lack of theoretical and empirical clarity ensures our science is not cumulative.

## Theories: Lost in Translation

Almost 30 years ago, Skinner identified over 100 terms used to describe the concept of 'control', and considered the theoretical and empirical challenges posed by such a preponderance of overlapping constructs (Skinner, 1996). Our ability to use and develop theory to predict, explain, and change behaviour requires we better attend to how constructs are identified, defined and measured (Benyamini et al., 2015; Dixon et al., 2024; Michell, 2020). A cumulative science requires that construct labels, including MoAs, and their definitions are explicit, shared, and agreed upon. The discipline needs, but still lacks, a method to agree labels and definitions for those labels, and how to modify both as new evidence emerges. Theoretical and empirical confusion is caused when different labels are used for the same construct and the same label used for different definitions (i.e., 'jingle/jangle') (Flake & Fried, 2020). Terminology is crucial when our measurement targets are not bounded by material constraints. It is also not enough to simply demonstrate that a measure is reliable (e.g., internally consistent) or to focus on common types of validity (e.g., predictive validity). Rather, our measures need to be able to discriminate one label and definition from another and to demonstrate that they do not measure other, related constructs, i.e., that they have both content and discriminant content validity (Bell et

al., 2017; Johnston et al., 2014).

## Organizing Theory: Mechanism Ontologies

The problematic lack of shared labels and definitions is not unique to the discipline of behavioural science. Other scientific fields face similar challenges, with different research groups using terms and measurements for constructs inconsistently (Larsen et al., 2017). Ontologies are classification frameworks that include representations of entities (anything that exists in the universe, e.g., objects and processes) with unique labels and definitions and relationships between these entities and can, therefore, address these barriers to scientific progress (Arp et al., 2015). Drawing on the success of ontologies in other disciplines (Gene Ontology Consortium, 2019), the MoA Ontology was developed to serve as a shared framework to communicate, operationalize, and synthesize evidence about MoAs for behavioural scientists (Michie et al., 2017; Schenk et al., 2023). Researchers can use this ontology to clearly operationalize MoAs when developing interventions, and to label and define MoAs more consistently when reporting research methods and results. Moreover, with its detailed entities, the ontology can be applied to synthesize evidence about MoAs, map measurements for MoAs, and identify research gaps for MoAs.

## From Theory to Practice: Measurement Matters

Once a mechanism has been clearly operationalized, it is necessary to find a valid, reliable, and sensitive measure. Although this may seem straightforward, there are myriad considerations that are often given insufficient

attention during study design, and measurement (and our science) have suffered as a result (Flake & Fried, 2020). First, the vast majority of hypothesized mechanisms are not observable. There is often no 'gold standard' or real-world benchmark by which we can judge the validity of our measures. Second, lack of shared vocabulary can lead to confusion regarding consistency of measurement across studies (Flake & Fried, 2020), which causes serious complications in evidence synthesis, among other issues. Third, measures may not be designed with the researcher-specified operationalization in mind; thus, measures may be incomplete (i.e., assess only a portion of the construct), may assess multiple mechanisms, or may assess a different (albeit related) construct entirely (Dixon & Johnston, 2019). For example, if a researcher hypothesizes that self-efficacy belief for a behaviour is the mechanism underlying an intervention effect, but the measure assesses beliefs about consequences, the hypothesis test is not valid.

Appropriate measurement of hypothesized mechanisms is essential for rigorous behavioural science. An ongoing collaboration between researchers at the Science Of Behavior Change (SOBC) program and the Human Behaviour Change Project (HBCP) aims to illuminate and address some of these issues through a project to identify links between 44 self-report measures, from the SOBC Measure Repository, to MoAs identified by the HBCP (Cornelius et al., 2023). Roundtable attendees were invited to participate in coding a selected measure to the MoA Ontology. The varied opinions highlighted lack of conceptual clarity in measure items and the difficulty of linking measures to precise and distinct mechanistic entities.

## Tools and Practical Considerations

Both SOBC and the HBCP have created scientific resources to facilitate mechanism-focused

behavioural science. The SOBC Measures Repository houses 140 measures (and counting), 114 of which represent putative mechanisms of behaviour change across three broad domains: self-regulation, stress and stress reactivity, and interpersonal processes (Science of Behaviour Change, [SOBC] 2021). Each measure includes Google scholar integration and download and scoring instructions. Measures from SOBC researchers additionally link to documentation on the Open Science Framework (osf). (Note that no assessment of measure quality was required for inclusion in the repository.)

The HBCP Theory and Techniques Tool (TaTT) summarizes evidence for (or against) links between 74 behaviour change techniques (BCTs) and 26 MoAs (i.e., might a given BCT influence a given MoA?) (Connel et al., 2019; Carey et al., 2019; Human Behaviour Change Project [HBCP], 2021). Additional work has updated the SOBC measures repository such that it now contains information about potentially related MoAs and links to the HBCP TaTT, and the TaTT updated such that each MoA includes links to potentially related measures. This greatly increases the interoperability of these resources, allowing researchers to navigate between the two tools as they form mechanistic hypotheses.

Additional considerations are critical in the selection, reporting, development, and evaluation of measurement instruments (Benyamini et al., 2015). Several indices of the quality of the measure are conventionally used. For self-report measures, it is common for internal consistency to be reported, but many other features need to be considered but are often overlooked. In an interactive demonstration, roundtable attendees were invited to think of a measure they had used or read about recently and evaluate it against 20 quality criteria, going beyond simple conventional assessments of reliability and validity. The quality criteria included: 1. *information about their measure prior to its use with respondents*, (e.g. adequacy of the construct definition, attributes to be measured, content and discriminant content

validity of items and response formats, readability for the intended respondents, rationale and consistency of scoring and aggregation); 2. *evaluation of data from respondents* (e.g. internal consistency and reliability, including, aspects of construct validity including structural aspects (e.g., factor analyses) and correlational or experimental evidence of the extent to which the measure predicts or is predicted as proposed by the framework, and sensitivity to change in the intended context). To conclude, attendees were asked to total the number of quality criteria met; only one measure, chosen by two delegates scored strongly. This exercise demonstrated that, although considerable work is being done to define and discriminate constructs, whether we are careful enough in measuring them remains a critical consideration.

## A Way Forward

As the title of the roundtable indicates, we are in double trouble with our measures. The jingle/jangle fallacy is akin to the Babylonian confusion of tongues, yet we tend to continue with business as usual, hoping that further data and statistical analyses will resolve our problems. Sophisticated statistical analyses can never replace proper, conceptual, and theoretical thinking. This roundtable is a wakeup call. First, it calls to be more precise in our definition of psychological concepts, which can be facilitated by clearly specified ontologies (Schenk et al., 2023). Second, it calls for a more precise measurement of our concepts (Cornelius et al., 2023; SOBC, 2021) and pleas for a re-evaluation of content validity. Content validity is a prerequisite for any other form of validity. However, it is often ignored and confused with face validity (i.e., the extent to which an instrument appears to be valid). Third, it calls for proper guidance in developing and selecting measures (as in the interactive

demonstration).

The problems addressed and solutions proposed in this roundtable are a way forward, but we must acknowledge and embrace key challenges. For example, validity is not a static property; measures are only valid for a particular purpose in a particular group in a particular setting (Van Ryckeghem & Crombez, 2022). Measures should always be carefully evaluated in terms of the question at hand. Furthermore, our concepts are not real or natural objects in the world or in our mind ('real entities'), they are theoretical constructs that work for us in the field of health psychology ('pragmatic entities'). Measures are then developed for a particular purpose. They are not a direct and fixed mapping of phenomena onto a numerical scale (Hand, 2016). Not surprisingly, different explanations and approaches are possible. As in biology and engineering, we may acknowledge and embrace ontological diversity (Ludwig, 2013). Methodological diversity should also be embraced. Bottom-up, qualitative approaches (cognitive interview) may complement top-down approaches (e.g., discriminant content validity) (Crombez et al., 2020; Horwood et al., 2010). We hope this roundtable will inspire and motivate behavioral scientists to pay more attention to defining and measuring constructs to improve scientific rigor and move the field forward.

## References

- Arp, R., Smith, B., & Spear, A.D. (2015). *Building ontologies with basic formal ontology*. MIT Press.
- Bell, C., Johnston, D., Allan, J., Pollard, B., & Johnston, M. (2017). What do Demand-Control and Effort-Reward work stress questionnaires really measure? A discriminant content validity study of relevance and representativeness of measures. *British journal of health psychology*, 22(2), 295–329. <https://doi.org/10.1111/bjhp.12232>
- Benyamini, Y., Johnston, M., & Karademas, E.C. (2015). *Assessment in Health Psychology*. Gottingen/Boston: Hogrefe.
- Carey, R. N., Connell, L. E., Johnston, M., Rothman, A. J., de Bruin, M., Kelly, M. P., & Michie, S. (2019). Behavior Change Techniques and Their Mechanisms of Action: A Synthesis of Links Described in Published Intervention Literature. *Annals of behavioral medicine : a publication of the Society of Behavioral Medicine*, 53(8), 693–707. <https://doi.org/10.1093/abm/kay078>
- Connell, L. E., Carey, R. N., de Bruin, M., Rothman, A. J., Johnston, M., Kelly, M. P., & Michie, S. (2019). Links Between Behavior Change Techniques and Mechanisms of Action: An Expert Consensus Study. *Annals of behavioral medicine: a publication of the Society of Behavioral Medicine*, 53(8), 708–720. <https://doi.org/10.1093/abm/kay082>
- Cornelius, T., Derby, L., Connell Bohlen, L., Birk, J. L., Rothman, A. J., Johnston, M., & Michie, S. (2023). Linking measures to mechanisms of action: An expert opinion study. *British journal of health psychology*, 28(1), 98–115. <https://doi.org/10.1111/bjhp.12614>
- Crombez, G., De Paepe, A. L., Veirman, E., Eccleston, C., Verleysen, G., & Van Ryckeghem, D. M. L. (2020). Let's talk about pain catastrophizing measures: an item content analysis. *PeerJ*, 8, e8643. <https://doi.org/10.7717/peerj.8643>
- Dixon, D., Den Daas, C., & Johnston, M.C. (2024). *Measurement in Health Psychology*. In K. Brown, C. Cheng, M. Hagger, K. Hamilton & S. Sutton (Eds.) *The Sage Handbook of Health Psychology*, SAGE Publications, 2nd edition.
- Dixon, D., & Johnston, M. (2019). Content validity of measures of theoretical constructs in health psychology: Discriminant content validity is needed. *British journal of health psychology*, 24(3), 477–484. <https://doi.org/10.1111/bjhp.12373>

- 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>
- Hand, D.J. (2016). *Measurement: A very short introduction*. Oxford University Press.
- Horwood, J., Pollard, B., Ayis, S., McIlvenna, T., & Johnston, M. (2010). Listening to patients: Verbal data in the pretesting of an Osteoarthritis health outcome instrument. *BMC Musculoskeletal Disorders*, 11, 182. <https://doi.org/10.1186/1471-2474-11-182>
- Human Behaviour Change Project (HBCP). (2021, June 28). *Theory and Techniques Tool. Human Behaviour Change Project*. <https://theoryandtechniquetool.humanbehaviourchange.org/>
- Johnston, M., Dixon, D., Hart, J., Glidewell, L., Schröder, C., & Pollard, B. (2014). Discriminant content validity: a quantitative methodology for assessing content of theory-based measures, with illustrative applications. *British journal of health psychology*, 19(2), 240–257. <https://doi.org/10.1111/bjhp.12095>
- Larsen, K. R., Michie, S., Hekler, E. B., Gibson, B., Spruijt-Metz, D., Ahern, D., Cole-Lewis, H., Ellis, R. J., Hesse, B., Moser, R. P., & Yi, J. (2017). Behavior change interventions: the potential of ontologies for advancing science and practice. *Journal of behavioral medicine*, 40(1), 6–22. <https://doi.org/10.1007/s10865-016-9768-0>
- Ludwig, David (2013). Disagreement in Scientific Ontologies. *Journal for General Philosophy of Science / Zeitschrift für Allgemeine Wissenschaftstheorie* (1):1-13.
- Michell, J. (2020). Thorndike's Credo: Metaphysics in psychometrics. *Theory & Psychology*, 30(3), 309-328. <https://doi.org/10.1177/0959354320916251>
- Michie, S., Thomas, J., Mac Aonghusa, P., West, R., Johnston, M., Kelly, M. P., Shawe-Taylor, J., Hastings, J., Bonin, F., & O'Mara-Eves, A. (2020). The Human Behaviour-Change Project: An artificial intelligence system to answer questions about changing behaviour. *Wellcome open research*, 5, 122. <https://doi.org/10.12688/wellcomeopenres.15900.1>
- Peters, G. Y., & Crutzen, R. (2022, November 22). *Knowing What We're Talking About: Facilitating Decentralized, Unequivocal Publication of and Reference to Psychological Construct Definitions and Instructions*. <https://doi.org/10.31234/osf.io/8tpcv>
- Schenk, P. M., Wright, A. J., West, R., Hastings, J., Lorencatto, F., Moore, C., Hayes, E., Schneider, V., & Michie, S. (2023). An ontology of mechanisms of action in behaviour change interventions. *Wellcome open research*, 8, 337. <https://doi.org/10.12688/wellcomeopenres.19489.1>
- Science Of Behavior Change (SOBC). (2012, June 28). The Measures. *Science Of Behavior Change*. <https://measures.scienceofbehaviorchange.org/>
- Skinner, E. A. (1996). A guide to constructs of control. *Journal of Personality and Social Psychology*, 71(3), 549–570. <https://doi.org/10.1037/0022-3514.71.3.549>
- Sumner, J. A., Beauchaine, T. P., & Nielsen, L. (2018a). A mechanism-focused approach to the science of behavior change: An introduction to the special issue. *Behaviour research and therapy*, 101, 1–2. <https://doi.org/10.1016/j.brat.2017.12.005>
- Sumner, J. A., Carey, R. N., Michie, S., Johnston, M., Edmondson, D., & Davidson, K. W. (2018b). Using Rigorous Methods to Advance Behaviour Change Science. *Nature human behaviour*, 2(11), 797–799. <https://doi.org/10.1038/s41562-018-0471-8>
- The Gene Ontology Consortium (2019). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic acids research*, 47(D1), D330–D338. <https://doi.org/10.1093/nar/gky1055>
- Van Ryckeghem, D. M. L., & Crombez, G. (2022). *Assessment and Measurement in Health Psychology*. In *Comprehensive Clinical*

Psychology, Second Edition (Vol. 8, pp. 85-94).  
Elsevier. <https://doi.org/10.1016/B978-0-12-818697-8.00193-X>



**Talea Cornelius**  
Center for Behavioral Cardiovascular  
Health, Columbia University Irving  
Medical Center  
[tmc2184@cumc.columbia.edu](mailto:tmc2184@cumc.columbia.edu)



**Paulina Schenk**  
Center of Behaviour Change,  
University College London  
[paulina.schenk.13@ucl.ac.uk](mailto:paulina.schenk.13@ucl.ac.uk)



**Diane Dixon**  
Edinburgh Napier University  
[D.Dixon@napier.ac.uk](mailto:D.Dixon@napier.ac.uk)



**Geert Crombez**  
Ghent University  
[Geert.Crombez@UGent.be](mailto:Geert.Crombez@UGent.be)



**Marie Johnston**  
University of Aberdeen, UK  
[m.johnston@abdn.ac.uk](mailto:m.johnston@abdn.ac.uk)