Four Challenges for Measurement in Environmental Psychology, and How to Address Them

Claudio D. Rosa [1], Eiko I. Fried [2], Lincoln R. Larson [3], and Silvia Collado [4]

[1] Department of Development and Environment, State University of Santa Cruz

[2] Department of Clinical Psychology, Leiden University

[3] Department of Parks, Recreation and Tourism Management, North Carolina State University

[4] Department of Psychology and Sociology, University of Saragossa

Author Note

Claudio D. Rosa http://orcid.org/0000-0002-1939-2716

Eiko I. Fried https://orcid.org/0000-0001-7469-594X

Lincoln R. Larson http://orcid.org/0000-0001-9591-1269

Silvia Collado https://orcid.org/0000-0002-3905-0617

We have no conflict of interest to disclose. This study was partially funded by the Spanish Ministry of Science, Education, and Universities (PGC2018-095502-B-I00)

and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Correspondence concerning this article should be addressed to Claudio D. Rosa, Department of Development and Environment, State University of Santa Cruz, Ilhéus-Bahia, Brazil. Email: claudio2008ilheus@hotmail.com. Phone: 55 73 991123063. Postal address: Avenida Governador Roberto Santos, number 289, Ilhéus/BA.

Four Challenges for Measurement in Environmental Psychology, and How to Address Them

Dear Editors:

Environmental psychologists are interested in a diverse array of constructs, including but not limited to environmental behaviors (Steg & Vlek, 2009; Urban & Braun Kohlová, 2022), environmental attitudes (Milfont & Duckitt, 2004; Wyss et al., 2022), environmental concern (Schultz, 2001), environmental beliefs (Dunlap & Van Liere, 1978; Rosa et al., 2022), and connection to nature (Coughlan et al., 2022; Ives et al., 2017). Valid and reliable measurement of such constructs is essential for research and practice. Invalid measurement can lead to misleading inferences. While most researchers are aware of this, recent work has revealed that many measures have critical limitations or are used inappropriately (Hawcroft & Milfont, 2010; Mokkink et al., 2018; Perrin & Benassi, 2009; Rosa et al., 2022; Stallwood et al., 2021; Terwee et al., 2018). In this letter, we use examples of frequently utilized and widely cited scales to illustrate four key challenges and corresponding recommendations regarding measurement in the field of environmental psychology. ~~We use the term "construct" to describe the concept or characteristic that a measure is designed to assess, such as observable behaviors or unobservable beliefs (AERA et al., 2014). We use multi-item measures as examples, but challenges are also relevant for single-item measures~~.

Clarify construct definition and operationalization. In psychological research, the construct(s) of interest should be ~~operationalized~~ defined in sufficient detail to ~~favor~~ enable the development of a measure that covers ~~only this~~ the particular target construct ~~and its full extension~~. The popular New Ecological Paradigm (NEP) offers one example of an ambiguously defined construct (Dunlap et al., 2000; Dunlap & Van Liere, 1978). Developers

of the NEP scale acknowledged uncertainty regarding what the scale was purported to measure, suggesting the NEP construct was "somewhat amorphous" (Dunlap et al., 2000, p. 429), including beliefs related to the balance of nature, the existence of ecological limits, and humans' role as a part of nature. It is difficult to know how well a scale assesses the construct it aims to measure without sufficient information about what that construct is (Rosa et al., 2022). In the case of the NEP, we cannot evaluate whether all dimensions of this amorphous construct are effectively covered by the items on the NEP scale(s). Whereas there is no consensus regarding what constitutes a sufficient characterization of a construct, it generally requires a definition as well as specific examples of what a construct is and what it is not ~~(Cortina et al., 2020; Flake & Fried, 2020)~~. Ideally, construct definition also considers theoretical relations with related constructs, for example, within a nomological network (Cortina et al., 2020; Flake & Fried, 2020). A further definition of the NEP might therefore include specifying all aspects of this construct and differentiating the NEP from other constructs like environmental attitudes, environmental beliefs, and connectedness to nature (Rosa et al., 2022).

Consider face validity and construct coverage. After describing the~~ir~~ construct(s) of interest in sufficient detail, researchers should ~~also~~ then show how the content of their measure matches the content of the construct. For example, the Connectedness to Nature Scale (CNS, Mayer & Frantz, 2004, p. 593) was designed to assess "individuals' trait levels of feeling emotionally connected to the natural world." Mayer and Frantz's (2004) description of the construct led other researchers in the field to question whether the CNS actually assesses feelings of emotional connection rather than related constructs such as beliefs about an individual's dependence on nature (Pasca et al., 2017; Perrin & Benassi, 2009). To investigate whether a measure matches the construct under investigation, scale developers

~~could~~ can gather expert opinions regarding the relevance of each scale item and~~/or~~ qualitatively describe how each item's content is related to the construct of interest (Terwee et al., 2018). For example, experts in connectedness to nature ~~could~~ can be provided with the construct definition and asked to indicate whether they think all the items of the CNS are relevant for this construct and if the items cover all aspects of connectedness to nature. Possible conclusions from such an evaluation could be that a construct needs to be better defined, that specific items are not related to the construct, or that items collectively fail to capture some important dimensions of the construct.

Examine interpretation of item(s) and response process in diverse participants. Valid measurement of a construct typically requires that a measure is interpreted by the target population as intended by the scale developers (AERA et al., 2014; Peterson et al., 2017). If respondents and scale developers differ in their interpretations, or if respondents differ from each other in their interpretations, or if a scale developer does not fully understand how respondents interpret items, measurement problems may ensue (AERA et al., 2014). Taking the NEP scale for children as an example, the item “People must still obey the laws of nature” is used to represent the idea of “human exemptionalism” within the broader NEP (Dunlap et al., 2000; Manoli et al., 2007). However, if children have a different understanding of the expression “laws of nature” than the one expected by the scale developers, the idea of “human exemptionalism” may be inaccurately represented (Harrison, 2020; Rosa et al., 2022). Similarly, varying conceptualizations of “nature” among children (Collado et al., 2016; Larson et al., 2011) and adults (Muhar et al., 2018) might lead researchers to draw inaccurate inferences related to items that include the word “nature.” Cognitive interviews, in which members of the target population respond to an instrument while expressing their thoughts aloud, can help researchers address this challenge through a better understanding of diverse

participants' response processes (Peterson et al., 2017). For example, Harrison (2020) interviewed children and adolescents to understand their thought processes when responding to items on the NEP scale for children, and Clayton et al. (2021) engaged in discussions and workshops across five countries to characterize potential cultural variations in participants' interpretation of the Environmental Identity Scale.

Align theoretical and statistical models. Scale developers nearly universally rely on reflective latent variable models (LVMs), such as exploratory and confirmatory factors models, to evaluate measures and justify changes (e.g., rephrasing or dropping items). These models impose ~~severe~~ assumptions on data that may require more consideration in the literature. One example is the common cause theory. In substantive language, this means that responses on items x_1, x_2, x_3,… x_n are caused (and only caused) by a common construct Y (Van Bork et al., 2017). For instance, because of the idea that the personality trait extraversion causes how people behave, this trait is measured with questions about certain types of behavior. Statistically, the common cause theory means that items are only correlated due to their association with the conditioning factor. Together, this implies that shared variance among items is due to the underlying construct, and variance that is not shared is measurement error (for details, see Fried, 2020; Rhemtulla et al., 2020; Van Bork et al., 2017). However, this assumption may not always be accurate, because different causal models can lead to shared variance among items x_1, x_2, x_3,… x_n other than a common factor Y; this might occur when one item x1 causes another item x2. For example, consider the following items from the Nisbet & Zelenski (2013) short version of the Nature Relatedness Scale (Y); "I feel very connected to all living things and the earth" (x1) and "My ideal vacation spot would be a remote, wilderness area" (x2). These two items may not be statistically independent given Y, because feeling connected with nature (x1) may stimulate

people's preferences for activities in remote natural areas (x2), which could in turn reinforce x1 (Barrable & Booth, 2020; Rosa et al., 2020). Therefore, we believe that researchers should provide an explanation as to why the LVM in particular is best suited for representing a focal construct at hand. Some authors have considered this. As an example, Kaiser and Lange (2021) propose that people's environmental attitudes can motivate engagement in environmental behaviors, so these attitudes should (at least in part) explain the shared variance among self-reported environmental behaviors.

These four measurement challenges are not unique to environmental psychology, and we do not see specific characteristics that make measurement harder in our field than in other areas of psychology. We chose to focus on these particular challenges because ~~(a)~~ they are very common in environmental psychology literature, ~~(b) they are relevant for both unobservable constructs and, in some cases, observable events and /behaviors,~~ and ~~(c)~~ they can, in principle, be resolved with greater attention to the content of ~~the~~ a construct and measure. Our aim in presenting these challenges is not to criticize decades of work in the field that has focused on the development of measures and assessment of constructs. Rather, we hope that outlining challenges and future directions will give rise to necessary discussions about measurement practices in environmental psychology that will help the field to better achieve its goals. We recommend that researchers consult other sources for further guidance on how to define a construct (Cortina et al., 2020; Flake & Fried, 2020; Rhemtulla et al., 2020), validate a measure (AERA et al., 2014; Cook et al., 2015; Kane, 2013; Mokkink et al., 2018; Terwee et al., 2018) and report a validation study (Flake & Fried, 2020; Gagnier et al., 2021). These sources offer guidance for addressing challenges not only during the development of new scales but also when utilizing previously developed scales. It is ultimately the researcher's responsibility to ensure that a chosen measure is valid for the

specific use to which it is being applied (AERA et al., 2014; Kane, 2013). Hopefully, our suggestions will help to improve the development and use of measures in environmental psychology and inspire future research on this important topic.

Declaration of Interest Statement

The authors declare they have no conflict of interest.

References

AERA, APA, & NCME. (2014). Standards for Educational and Psychological Testing. In *American Educational Research Association (AERA), American Psychological Association (APA), National Council on Measurement in Education (NCME)*. American Educational Research Association.

Barrable, A., & Booth, D. (2020). Increasing nature connection in children: A mini review of interventions. *Frontiers in Psychology*, *11*(492), 1–7. https://doi.org/10.3389/fpsyg.2020.00492

Clayton, S., Czellar, S., Nartova-Bochaver, S., Skibins, J. C., Salazar, G., Tseng, Y.-C., Irkhin, B., & Monge-Rodriguez, F. S. (2021). Cross-cultural validation of a revised environmental identity scale. *Sustainability*, *13*(4), 2387. https://doi.org/10.3390/su13042387

Collado, S., Íñiguez-Rueda, L., & Corraliza, J. A. (2016). Experiencing nature and children's conceptualizations of the natural world. *Children's Geographies*, *14*(6), 716–730. https://doi.org/10.1080/14733285.2016.1190812

Cook, D. A., Brydges, R., Ginsburg, S., & Hatala, R. (2015). A contemporary approach to validity arguments: A practical guide to Kane's framework. *Medical Education*, *49*(6), 560–575. https://doi.org/10.1111/medu.12678

Cortina, J. M., Sheng, Z., Keener, S. K., Keeler, K. R., Grubb, L. K., Schmitt, N., Tonidandel, S., Summerville, K. M., Heggestad, E. D., & Banks, G. C. (2020). From alpha to omega and beyond! A look at the past, present, and (possible) future of psychometric soundness in the Journal of Applied Psychology. *Journal of Applied Psychology*. https://doi.org/10.1037/apl0000815

Coughlan, A., Ross, E., Nikles, D., De Cesare, E., Tran, C., & Pensini, P. (2022). Nature

guided imagery: An intervention to increase connectedness to nature. *Journal of Environmental Psychology*, *80*, 101759. https://doi.org/10.1016/j.jenvp.2022.101759

Dunlap, R. E., & Van Liere, K. D. (1978). The "New Environmental Paradigm": A Proposed measuring instrument and preliminary results. *The Journal of Environmental Education*, *9*(4), 10–19. https://doi.org/10.1080/00958964.1978.10801875

Dunlap, R. E., Van Liere, K. D., Mertig, A. G., & Jones, R. E. (2000). New trends in measuring environmental attitudes: Measuring endorsement of the New Ecological Paradigm: A Revised NEP Scale. *Journal of Social Issues*, *56*(3), 425–442. https://doi.org/10.1111/0022-4537.00176

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

Fried, E. I. (2020). Lack of theory building and testing impedes progress in the factor and network literature. *Psychological Inquiry*, *31*(4), 271–288. https://doi.org/10.1080/1047840X.2020.1853461

Gagnier, J. J., Lai, J., Mokkink, L. B., & Terwee, C. B. (2021). COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Quality of Life Research*, *30*(8), 2197–2218. https://doi.org/10.1007/s11136-021-02822-4

Harrison, G. M. (2020). Validity evidence against the children's New Ecological Paradigm Scale. *The Journal of Environmental Education*, *51*(1), 1–13. https://doi.org/10.1080/00958964.2019.1646202

Hawcroft, L. J., & Milfont, T. L. (2010). The use (and abuse) of the new environmental paradigm scale over the last 30 years: A meta-analysis. *Journal of Environmental Psychology*, *30*(2), 143–158. https://doi.org/10.1016/j.jenvp.2009.10.003

Ives, C. D., Giusti, M., Fischer, J., Abson, D. J., Klaniecki, K., Dorninger, C., Laudan, J., Barthel, S., Abernethy, P., Martín-López, B., Raymond, C. M., Kendal, D., & von Wehrden, H. (2017). Human–nature connection: A multidisciplinary review. *Current Opinion in Environmental Sustainability*, *26–27*, 106–113. https://doi.org/10.1016/j.cosust.2017.05.005

Kaiser, F. G., & Lange, F. (2021). Offsetting behavioral costs with personal attitude: Identifying the psychological essence of an environmental attitude measure. *Journal of Environmental Psychology*, *75*(101619), 1–10. https://doi.org/10.1016/j.jenvp.2021.101619

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, *50*(1), 1–73. https://doi.org/10.1111/jedm.12000

Larson, L. R., Green, G. T., & Castleberry, S. B. (2011). Construction and validation of an instrument to measure environmental orientations in a diverse group of children. *Environment and Behavior*, *43*(1), 72–89. https://doi.org/10.1177/0013916509345212

Manoli, C. C., Johnson, B., & Dunlap, R. E. (2007). Assessing children's environmental worldviews: Modifying and validating the New Ecological Paradigm scale for use with children. *The Journal of Environmental Education*, *38*(4), 3–13. https://doi.org/10.3200/JOEE.38.4.3-13

Mayer, F. S., & Frantz, C. M. (2004). The Connectedness to Nature Scale: A measure of individuals' feeling in community with nature. *Journal of Environmental Psychology*, *24*(4), 503–515. https://doi.org/10.1016/j.jenvp.2004.10.001

Milfont, T. L., & Duckitt, J. (2004). The structure of environmental attitudes: A first- and second-order confirmatory factor analysis. *Journal of Environmental Psychology*, *24*(3), 289–303. https://doi.org/10.1016/j.jenvp.2004.09.001

Mokkink, L. B., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Terwee, C. B. (2018). *COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) - user manual*. COSMIN. https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf

Muhar, A., Raymond, C. M., van den Born, R. J. G., Bauer, N., Böck, K., Braito, M., Buijs, A., Flint, C., de Groot, W. T., Ives, C. D., Mitrofanenko, T., Plieninger, T., Tucker, C., & van Riper, C. J. (2018). A model integrating social-cultural concepts of nature into frameworks of interaction between social and natural systems. *Journal of Environmental Planning and Management*, *61*(5–6), 756–777. https://doi.org/10.1080/09640568.2017.1327424

Nisbet, E. K., & Zelenski, J. M. (2013). The NR-6: A new brief measure of nature relatedness. *Frontiers in Psychology*, *4*(813). https://doi.org/10.3389/fpsyg.2013.00813

Pasca, L., Aragonés, J. I., & Coello, M. T. (2017). An analysis of the Connectedness to Nature Scale based on item response theory. *Frontiers in Psychology*, *8*(AUG). https://doi.org/10.3389/fpsyg.2017.01330

Perrin, J. L., & Benassi, V. A. (2009). The Connectedness to Nature Scale: A measure of emotional connection to nature? *Journal of Environmental Psychology*, *29*(4), 434–440. https://doi.org/10.1016/j.jenvp.2009.03.003

Peterson, C. H., Peterson, N. A., & Powell, K. G. (2017). Cognitive interviewing for item development: Validity evidence based on content and response processes. *Measurement and Evaluation in Counseling and Development*, *50*(4), 217–223. https://doi.org/10.1080/07481756.2017.1339564

Rhemtulla, M., van Bork, R., & Borsboom, D. (2020). Worse than measurement error:

Consequences of inappropriate latent variable measurement models. *Psychological Methods*, *25*(1), 30–45. https://doi.org/10.1037/met0000220

Rosa, C. D., Collado, S., & Larson, L. R. (2022). The utility and limitations of the New Ecological Paradigm scale for children. *The Journal of Environmental Education*, 1–12. https://doi.org/10.1080/00958964.2022.2044281

Rosa, C. D., Larson, L. R., Collado, S., Cloutier, S., & Profice, C. C. (2020). Gender differences in connection to nature, outdoor preferences, and nature-based recreation among college students in Brazil and the United States. *Leisure Sciences*, 1–21. https://doi.org/10.1080/01490400.2020.1800538

Schultz, P. W. (2001). The structure of environmental concern: Concern for self, other people, and the biosphere. *Journal of Environmental Psychology*, *21*(4), 327–339. https://doi.org/10.1006/jevp.2001.0227

Stallwood, E., Monsour, A., Rodrigues, C., Monga, S., Terwee, C., Offringa, M., & Butcher, N. J. (2021). Systematic review: The measurement properties of the Children's Depression Rating Scale−Revised in adolescents with major depressive disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *60*(1), 119–133. https://doi.org/10.1016/j.jaac.2020.10.009

Steg, L., & Vlek, C. (2009). Encouraging pro-environmental behaviour: An integrative review and research agenda. *Journal of Environmental Psychology*, *29*(3), 309–317. https://doi.org/10.1016/j.jenvp.2008.10.004

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., de Vet, H. C. W., Bouter, L. M., Alonso, J., Westerman, M. J., Patrick, D. L., & Mokkink, L. B. (2018). *COSMIN methodology for assessing the content validity of PROMs - user manual*. COSMIN. https://www.cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-

user-manual-v1.pdf

Urban, J., & Braun Kohlová, M. (2022). The COVID-19 crisis does not diminish environmental motivation: Evidence from two panel studies of decision making and self-reported pro-environmental behavior. *Journal of Environmental Psychology*, *80*, 101761. https://doi.org/10.1016/j.jenvp.2022.101761

Van Bork, R., Wijsen, L. D., & Rhemtulla, M. (2017). Toward a causal interpretation of the common factor model. *Disputatio*, *9*(47), 581–601. https://doi.org/10.1515/disp-2017-0019

Wyss, A. M., Knoch, D., & Berger, S. (2022). When and how pro-environmental attitudes turn into behavior: The role of costs, benefits, and self-control. *Journal of Environmental Psychology*, *79*, 101748. https://doi.org/10.1016/j.jenvp.2021.101748