Methodological Advances in Motor Learning and Development

Keith R. Lohse

University of Utah

Motor learning and development are some of the most interesting fields in which we can exhaust ourselves with scientific inquiry. For one thing, the applications of these fields are so wide ranging: helping children master the foundations of physical activity, training doctors to perform delicate surgical procedures, pushing elite athletes to new heights of performance, or helping a client re-master lost movement following a brain injury. At times, our research may be far removed from these worldly applications, but it is important to understand the fundamentals of how the individual and their environment interact to determine motor functioning (Newell, 1991; World Health Organization, 2001).

We also need to appreciate the diversity of approaches that is required for putting that research into practice (Woolf, 2008). Basic research shapes our theoretical understanding and establishes the ideal *efficacy* of interventions, while applied research translates and implements these principles to determine their *effectiveness* outside of the laboratory (Brook & Lohr, 1985; Singal, Higgins, & Waljee, 2014). Naturally, we will each focus our respective works more at one level of analysis than another (e.g., I study motor behavior more than neural control of movement), but we should all appreciate the balance between different levels of analysis (Poggio, 2012). As our research questions change, so do our measures, methods, and theories. Neurophysiology is important, but psychology is not just ‘applied’ physiology, nor is coaching just ‘applied’ motor learning (Anderson, 1972). As Douglas Adams quipped, “If you try to take a cat apart to see how it works, the first thing you have on your hands is a non-working cat” (Adams, 2002).

Tempering my excitement about studying motor learning and development, I must admit we also live in an interesting time to be doing research. Psychological and biomedical research are said to be in a “replication crisis” (e.g., Open Science Collaboration, 2015; Patil, Peng, & Leek, 2016; Prinz, Schlange, & Asadullah, 2011), which has also been more positively dubbed a “credibility revolution” (Vazire, 2017). New terms like “p-hacking” (Simmons, Nelson, & Simonsohn,

Lohse (rehabinformatics@gmail.com) is with the Department of Health, Kinesiology, and Recreation; and the Department of Physical Therapy and Athletic Training; University of Utah, Salt Lake City, UT, USA.

2011), “HARKing” (hypothesizing after results are known; Kerr, 1998), and “QRPs” (questionable research/reporting practices; John, Loewenstein, & Prelec, 2012; Wigboldus & Dotsch, 2016) have all entered the scientific lexicon. At the same time, statistical and philosophical debates about inference are becoming more mainstream (Benjamin et al., 2018; McShane, Gal, Gelman, Robert, & Tackett, 2019). The *American Statistical Association* (ASA hereafter) has even gone on record to discourage the use of the term “statistical significance” (Wasserstein, Schirm, & Lazar, 2019). In fairness to the ASA, this is largely because, as applied researchers, we have repeatedly gotten the definition of p -values wrong and have not used the concept of significance wisely (Nuzzo, 2015; Wasserstein & Lazar, 2016).

These debates can feel tumultuous and have led many researchers to reflect (and perhaps despair) on their methodological practices. There is a lot of debate about how to solve these problems—and that debate only occurs when researchers can even agree on what the problems really are. One very positive solution, I think, is that open science is becoming increasingly standard. At least within certain definitions of “open”; great strides remain for diversity and representation in science (Freeman, 2018; Lallensack, 2017; Yoder & Mattheis, 2016).

With respect to methodological openness, authors have pushed for greater transparency, pre-specified analysis plans, and sharing data to the extent it is appropriate (Bishop, 2019; Munafò et al., 2017). An increasing number of journals also promote data sharing, the archiving of pre-prints, and the publication of Registered Reports to reduce publication bias (Chambers, Dienes, McIntosh, Rotshtein, & Willmes, 2015). However, much like the ASA recommends the thoughtful interpretation of p -values, we must approach these open science advances carefully as well. For instance, replications are important but sadly undervalued in many areas of science (Makel, Plucker, & Hegarty, 2012). I would emphatically argue that we need more direct replications. However, I would also argue that we need a healthy balance of epistemic approaches in science (Devezer, Nardin, Baumgaertner, & Buzbas, 2019). Yes, we want to know which effects are replicable and which effects are not, but this does not mean that we should shy away from attempting irreproducible research, provided we also attempt the replications.

Similarly, I think that pre-registration is an invaluable tool for increasing transparency (Nosek et al., 2019) and that adequate sample sizes are essential for replicable results (Button et al., 2013; Lohse, Buchanan, & Miller, 2016). However, it is also important that we do not mistake these practices as markers of improved scientific reasoning (Szollosi et al., 2019). To draw on a historical example, early physical investigations into the nature of temperature were highly replicable, but conceptualizing heat as a “caloric fluid” was still incorrect (Chang, 2004). Indeed, it was the falsification of theory and clever experimental design (in equally replicable experiments) that got us to the “kinetic theory” of temperature. As such, open research practices facilitate but are not alone sufficient for advancing our knowledge. Yes, we should absolutely make sure our findings are reliable before we weave theories to explain them. Yes, we should be as transparent as possible so that the relative strengths and weaknesses of our approaches can be debated by experts. Even then though, simply knowing the conditions for reliably eliciting an effect does not mean we understand “why” an effect occurs.

It is a tad depressing to accept that we may never get to ultimate causes, but as a famous *bon mot* asserts, “all models are wrong, some are useful”. In science as in life, I’d like to be less wrong about things as often and as quickly as possible. (Rest assured this is not merely rhetoric, but a genuine aspiration.)

These topics are what make me excited to introduce this special issue of the *Journal of Motor Learning and Development*, “Methodological Advances in Motor Learning and Development”. As substantive knowledge in a field moves forward, so do the methods that we have available for asking scientific questions. In this special issue, there are authors from a range of research areas who have made advances in research methods. I think that all these articles will give motor learning and development researchers something to think about, and potentially change how they approach their own research. Below, I break the discussion of methodology into constituent pieces of theory, design, measurement, and analysis. All these areas are critical to scientific progress and I will not give any of them their full credit here. In each section, I will try to briefly explain why I think each area is important and touch on major findings/arguments made by the articles in this special issue. Please keep in mind that this is an editorial, so below I will be heavily editorializing.

Theory

Although the precise definition of a scientific theory is debated, a “theory” can broadly be thought to be a conceptual model of the natural world that can be directly tested with empirical observation (Cartwright, 1997; Chalmers, 2013). Theories need not be experimentally tested (e.g., many astronomical or geological theories), but often we will seek to falsify a theory’s predictions through carefully controlled experimental conditions. Theory is also the first step in any methodological question, informing the design, selection of measures, and method of analysis. Colloquially, theories are often described as “weak” versus “strong”. More formally, we should describe a theory as making weak versus strong predictions (Fiedler, 2017; Meehl, 1967). Theoretical strength is a bit ambiguous, but encompasses notions of completeness (e.g., unified theories) and specificity (e.g., qualitative vs. quantitative predictions). Within motor learning and development, I would argue that we have weak theoretical foundations. We have identified many key constructs through years of important investigation, but we have yet to bring these constructs together into a cohesive ontology. We have many seminal theoretical works in our field (e.g., Adams, 1971; Kawato, 1999; Latash, 2010; Schmidt, 1975; Thelen & Smith, 1996; Wulf & Lewthwaite, 2016), but at the best of times our theories tend towards qualitative predictions rather than quantitative predictions. This weakness is not unusual given the young age of our field. I mean no insult to many years of fantastic work when I say that I think our field has a lot of important findings, but weak theory.

An exhaustive discussion of what constitutes strong or weak theory is beyond the scope of my introduction, but I’ve tried to adopt a heuristic approach as explained in Box 1. By scrutinizing the degree to which theory determines study design, we can say whether a given theory makes strong predictions. Admittedly, this is just a heuristic, but I think that most designs in motor learning and

development would fall to the left-hand side of Figure 1 and therefore be weak theories. Turning again to history as an illustrative example, when physicists thought they had solved the problem of cold fusion (see Ball, 2019 for discussion), other teams were able to quickly mobilize with agreement on how to test these claims. The speed and consensus on how to test these claims came in part because strong theory determined exactly *how* these claims should be tested. The potentially life changing announcement was realized to be an illusory finding in less than four months!

In contrast, “ego-depletion” was first introduced in psychology in the late 1990s (Baumeister, Bratslavsky, Muraven & Tice, 1998), although related

Box 1: Do I Have a Strong Theory? A Heuristic Approach

I argue that the degree to which a theory can be said to be “weak” or “strong” can be approximated from how researchers approached the study design. All studies have design parameters that the researchers must select:

- Who to include in the sample?
- How to administer the intervention and in what “dose”?
- What specific dependent measures to collect and how often?
- What is the relative magnitude of your “effect”?
- What factors need to be controlled for (either experimentally or statistically)?
- What control conditions (either positive or negative) need to be included? (and so on)

The more these parameters are free to vary, the weaker your theory is, because the theory is not effectively ruling out any design parameters. Again, there are no cut-offs for when a theory transitions from weak to strong, but Figure 1 shows some expressions that researchers might use when designing a study. For each design parameter, you can think about where your response would fall from left to right in Figure 1. If most of your responses fall on the left, your theory makes weak predictions. If most of your responses fall on the right, your theory makes strong predictions.

Keep in mind, having a theory make strong predictions does not mean that your theory is correct. Your theory could make strong predictions and still be wrong (like the “caloric fluid” theory of temperature; Chang, 2004). Stronger theories are, however, much more testable (or falsifiable if you prefer; Popper, 2005). Having a strong theory is thus desirable because you don’t have to fret about “Did I design the study correctly?” You can instead reject the theory (at least in its current form) and go on about the business of being less wrong.

Figure 1 — If you find yourself saying things on the left-hand side, you likely have a weak theory. The more you say things on the right-hand side, you have a stronger theory that makes strong predictions. If you are saying things in middle, your theory is moderate (Z is ruled out as a design parameter) but has room to improve (Is design parameter X or Y more appropriate?).

constructs are older. Almost twenty years later, meta-analyses (Carter & McCullough, 2014) and multi-lab replication efforts (Hagger et al., 2016) strongly suggested there was no compelling evidence for ego-depletion. Far from dead though, ego-depletion lives on in counterarguments (Cunningham & Baumeister, 2016) or in re-conceptualizations (Lin, Saunders, Friese, Evans, & Inzlicht, 2019). Naturally there are a lot of factors at play in this protracted debate (including sociological ones), but part of the problem stems from weak psychological theories that allow researchers to exploit degrees of freedom in their study designs (Simmons et al., 2011). To be clear, I am not psychologist enough to say if ego-depletion is real or if these debates are a necessary stage of larval empiricism before the debate pupates into strong theory. However, I do think that weak theory contributed to the lengthy timeline of these debates. The design parameters left open by weak theory allow a “research question” to become amorphous, and exploiting those degrees of freedom allows for many different possible patterns of data to support an ill-defined theory.

A motor learning example of which I am guilty is hypothesizing that changing practice conditions (e.g., self-control of difficulty) will improve skill “learning”. In my experimental design, however, “learning” might be operationalized in several different ways. By improved learning, do I mean both delayed retention and transfer tests? Do I mean just one of those tests? Or, do I even mean that I am willing to totally conflate two different constructs and call an increased rate of acquisition improved learning? (For a distinction see Kantak & Winstein, 2012). Again, if the operational definition of learning is left entirely up to me as the researcher, I have a weak theory. Naturally, it would be great to have a strong theory that reduces these degrees of freedom, but theory development takes time. This is why I think transparency of reporting is important; while we work to improve our theories, we need to let our readers know exactly what we planned to do and what we ultimately did do (Nosek et al., 2019; Vazire, 2017).

Given my pessimistic view of theory, it probably won’t surprise you that I would also say none of the articles in this special issue include theoretical revelations. But I also do not think this is a bad thing. The articles in this issue make many other important advancements that will improve the research process. The methodological advancements described later in this editorial will help us to test new hypotheses (Hauge et al., 2020; Hooyma, Garbin, & Fisher, 2020), adopt more replicable approaches to analysis so that research findings are more reliable (Chagdes, Liddy, Arnold, Claxton, & Haddad, 2020; Immekus, Muntis, & Terson de Paleville, 2020), or encourage us to think about constructs in new ways (Garcia-Masso, Estevan, Izquierdo-Herrera, Villarrasa-Sapiña, & Gonzalez, 2020; Lucca, Gire, Horton, & Somerville, 2020; Ng, Button, Collins, Giblin, & Kennedy, 2020; Wang, Infurna, & Schaefer, 2020).

In time, I think these methodological advancements will help us develop stronger theoretical predictions. In reflecting on the content of the special issue, however, I felt there was a lack of theoretical development that was surpassed by methodological development in other areas. I think researchers should see theory as an important frontier of methodological development. Currently though, I worry we have a plethora of findings in search of strong predictive models (Forscher, 1963).

Design

A common aspect of design in motor learning and development is that we are dealing with observations over time. Motor development often deals with time-scales of months, years or even decades, whereas motor learning often deals with shorter time-scales of trials, days, or weeks. As such, motor development tends to include more cross-sectional designs than motor learning (e.g., comparing children of different ages), but both disciplines are focused on explaining change over time.

Yet design encompasses more than which factors are crossed or nested in a given study. For instance, it is common that we might design a study to test for a difference between groups, but we might also be interested in testing the equivalence between two groups. Additionally, once a manipulation is chosen, we have considerable flexibility in how that intervention is delivered (e.g., the relative frequency, intensity, or timing). In this special issue, there are papers from several research teams who make significant advances to study design in motor learning and development (Chagdes et al., 2020; Fears & Lockman, 2020; Hooyman et al., 2020; Wang et al., 2020). These articles make advances in other areas as well, but upon reading, their major contributions struck me as study design.

Chagdes et al. (2020) discuss some of the challenges and technical limitations of using portable force measurement devices outside of the laboratory. Using mathematical models, the authors show how center of pressure errors can be accounted for by the posture being examined, the anthropometrics of the individual, and the nature of the support surface. Accounting for these errors opens new opportunities for developmental researchers to measure center of pressure data outside the laboratory.

Fears and Lockman (2020) present a perception-action account of the relationship between eye-movements and handwriting in children. They further describe how head-mounted eye-tracking technologies can be used to measure children's eye-movements in real time to better understand the development of handwriting.

Hooyman et al. (2020) present a perspective paper on a novel method for pair-associative stimulation using transcranial magnetic stimulation. Including some feasibility data, the authors propose how this method could be used to modulate intra-cortical connectivity and discuss its applications in motor learning.

Wang et al. (2020) adopt a framework for separately modeling within-session changes and long-term retention. In cognitively intact older adults, visuospatial ability was specifically associated with retention 24-hrs later. These effects were specific to visuospatial ability as no other subscales of the Montreal Cognitive Assessment were reliably associated and specific to long-term learning, as within-session changes were not reliably associated either.

Measurement

Measurement is fundamental to any scientific inquiry. Indeed, without it, the other factors here would matter very little (Flake & Fried, 2019). For example, if we set out to measure children's physical health as a function of their motor competence

and parental climate, and we don't have *valid* measures of those constructs, then it doesn't matter how big our sample size, nor how beautiful our theorizing.

Motor learning and development are also in the interesting position of combining many *direct* measures (e.g., force, speed, time) and *indirect* measures (e.g., psychological characteristics, multi-dimensional variables like "impairment"). It is tempting to think that the signal-to-noise ratio is greater for physical measures, or that reliability and validity are less of a concern, but that is not true (e.g., [Bennett, Miller, & Wolford, 2009](#); [Hedges, 1987](#)). Considerations about internal validity, external validity, statistical conclusion validity, and construct validity apply to all types of measures. In this special issue, several teams of researchers have made significant advances in the area of measurement ([Field, Esposito Bosma, & Temple, 2020](#); [Garcia-Masso et al., 2020](#); [Hauge et al., 2020](#); [Lucca et al., 2020](#); [Ng et al., 2020](#)). As before, these papers make several contributions, but I perceive their primary contribution to be in measurement.

Garcia-Masso et al. (2020) show how researchers measure postural control in children using portable force measurement devices and a "self-organizing maps" analysis. This data-driven method searches for clusters of similar individuals across the multi-dimensional space of different postural variables. Further, the authors show how these latent clusters relate to children's age, sex, and height.

Hauge et al. (2020) demonstrate the utility of Levenshtein distance as a measure of high-level motor planning when comparing two sequences. In information theory, the Levenshtein distance is computed as the "distance" between two strings in terms of the smallest number of substitutions, deletions, and insertions required to translate between the two. For instance, the distance between "cat" and "hat" is one, because "h" needs only be substituted for "c", whereas the distance between "cat" and "hatch" is three. Adapting this metric to motor control offers researchers a method for measuring the complexity of high-level motor plans.

Lucca et al. (2020) present a systematic review of the methods that have been traditionally used to study persistence behavior in infants. Using a case study, they also argue that measures of force and motion, common in the motor domain, may offer insights to infants' persistence. Used in conjunction with other measures, it is exciting to think about how motor behavior may offer us new insights into research on personality and temperament.

Ng et al. (2020) examined the reliability and construct validity of a new tool for assessing motor competence using the Microsoft Kinect[®]. In a sample of $N = 83$ children, their General Movement Competence Assessment showed promising internal validity and identified four movement constructs: locomotion, object control, stability, and a fourth new construct the authors tentatively identify as dexterity. These initial results are very interesting, especially given the automated nature of the measurement tool. With further research, the General Movement Competence Assessment may provide researchers with a low-cost tool for measuring motor competence in children.

Field et al. (2020) compare second and third editions of the Test of Gross Motor Development in a sample of $N = 270$ children from Grade 3 to Grade 5. Although correlated, scores on the third edition of the test tended to be slightly lower. There was also evidence for a consistent pattern for sex differences in the object control/ball skill sections of these tests. These differences suggest not only the need for sex-based norms, but also call into question why these differences

emerged. I am not a development researcher, but to me this suggests not bias in the assessment, but bias in the life experiences of young boys and girls that is reflected in a valid test.

Analysis

Statistical inferences and p -values, in particular, receive a disproportionate amount of attention in the popular discussion of statistical analysis (Leek & Peng, 2015). Indeed, data cleaning, exploratory data analysis, dealing with missing data, and the choice of statistical model are all design parameters that researchers need to consider. Most critically, these choices should be driven by the research question, so it is important that researchers appreciate the diversity of models available. If we do the reverse, and shoehorn our research questions into known statistical models, we risk designing sub-optimal studies or worse. For instance, many studies have attempted to answer questions of within-subject associations by making inferences from between-subject designs or other similar failures to account for how data are clustered (e.g., “Simpson’s Paradox”, Snijders & Boskers, 2012). In this special issue, research teams showcase novel applications of analyses or make arguments for including more types of analyses in motor learning and development (Immekus et al., 2020; Lohse, Shen, & Kozlowski, 2020; Mitchell, Loovis, & Butterfield, 2020).

Both Mitchell et al. (2020) and Lohse et al. (2020) point out that although motor learning and development researchers are regularly dealing with time-series data, most researchers rely on repeated measures ANOVA as a method of analysis and seldom adopt other methods for analyzing longitudinal data. Although Mitchell et al. use the term “hierarchical linear models” and Lohse et al. more generally argue for “mixed-effect regression” models, the two arguments are intermingled. Across the two articles, readers will gain a valuable understanding of advanced methods for longitudinal data analysis, some of the limitations of repeated measures ANOVA, and consider which analytical tools are most appropriate for their research questions.

Immekus et al. (2020) similarly demonstrate advances in regression analysis in a study of the association between children’s academic performance with motor proficiency. Specifically, these authors provide an illustrative example of a least absolute shrinkage and selection operator (or LASSO) in their regression analyses. “Shrinkage” refers to the fact that ordinary least squares estimates are going to provide the strongest correspondence to the sample data that generate their estimates. Thus, on average, our r^2 values or parameter estimates will tend to “shrink” relative to the original model, when the same model is fit in a new sample. LASSO and other similar methods provide useful tools for improving the generalization of our results from sample to sample.

Conclusions

I am honored to introduce this special issue of the *Journal of Motor Learning and Development*. Through my meandering introduction, I hope I have impressed upon you: what I perceive as the importance of open science; what I perceive as a lack of

strong theory in our research; and the numerous advances that articles in this special issue make to design, measurement, and analysis. I hope the issue is as fun to read as it was to help put together.

Acknowledgments

I would like to thank the many authors, handling editors, and anonymous reviewers who made this special issue possible. I would also like to thank Drs. Daniela Corbetta, Matt Miller, and Crissa Levin for providing feedback on early drafts of this editorial. The author received no funding specifically to pursue this work.

References

- Adams, D. (2002). *The salmon of doubt: Hitchhiking the universe one last time*. New York, NY: Harmony Books.
- Adams, J.A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior*, 3(2), 111–150. PubMed ID: [15155169](#) doi:[10.1080/00222895.1971.10734898](#)
- Anderson, P.W. (1972). More is different. *Science*, 177(4047), 393–396. PubMed ID: [17796623](#) doi:[10.1126/science.177.4047.393](#)
- Ball, P. (2019). Lessons from cold fusion, 30 years on. *Nature*, 569(7758), 601–601 PubMed ID: [31133704](#) doi:[10.1038/d41586-019-01673-x](#)
- Baumeister, R.F., Bratslavsky, E., Muraven, M., & Tice, D.M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74(5), 1252–1265. PubMed ID: [9599441](#) doi:[10.1037/0022-3514.74.5.1252](#)
- Benjamin, D.J., Berger, J.O., Johannesson, M., Nosek, B.A., Wagenmakers, E.J., Berk, R., . . . Cesarini, D. (2018). Redefine statistical significance. *Nature Human Behaviour*, 2(1), 6–10. PubMed ID: [30980045](#) doi:[10.1038/s41562-017-0189-z](#)
- Bennett, C.M., Miller, M.B., & Wolford, G.L. (2009). Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction. *Neuroimage*, 47(Suppl. 1), S125. doi:[10.1016/S1053-8119\(09\)71202-9](#)
- Bishop, D. (2019). Rein in the four horsemen of irreproducibility. *Nature*, 568(7753), 435–435. PubMed ID: [31019328](#) doi:[10.1038/d41586-019-01307-2](#)
- Brook, R.H., & Lohr, K.N. (1985). Efficacy, effectiveness, variations, and quality: Boundary-crossing research. *Medical Care*, 23(5), 710–722. PubMed ID: [3892183](#) doi:[10.1097/00005650-198505000-00030](#)
- Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., & Munafò, M.R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376. PubMed ID: [23571845](#) doi:[10.1038/nrn3475](#)
- Carter, E.C., & McCullough, M.E. (2014). Publication bias and the limited strength model of self-control: Has the evidence for ego depletion been overestimated? *Frontiers in Psychology*, 5, 823. PubMed ID: [25126083](#)
- Cartwright, N. (1997). Models: The blueprints for laws. *Philosophy of Science*, 64, S292–S303. doi:[10.1086/392608](#)
- Chagdes, J.R., Liddy, J.J., Arnold, A.J., Claxton, L.J., & Haddad, J.M. (2020). A mathematical model to examine issues associated with using portable force-measurement technologies to collect infant postural data. *Journal of Motor Learning and Development*, 8, 14–37. doi:[10.1123/jmld.2019-0009](#)

- Chalmers, A.F. (2013). *What is this thing called science?* Indianapolis, IN: Hackett Publishing.
- Chambers, C.D., Dienes, Z., McIntosh, R.D., Rotshtein, P., & Willmes, K. (2015). Registered reports: Realigning incentives in scientific publishing. *Cortex*, 66, A1–A2. PubMed ID: [25892410](#) doi:[10.1016/j.cortex.2015.03.022](#)
- Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford, UK: Oxford University Press.
- Cunningham, M.R., & Baumeister, R.F. (2016). How to make nothing out of something: Analyses of the impact of study sampling and statistical interpretation in misleading meta-analytic conclusions. *Frontiers in Psychology*, 7, 1639. PubMed ID: [27826272](#) doi:[10.3389/fpsyg.2016.01639](#)
- Devezer, B., Nardin, L.G., Baumgaertner, B., & Buzbas, E.O. (2019). Scientific discovery in a model-centric framework: Reproducibility, innovation, and epistemic diversity. *PLoS One*, 14(5), e0216125. PubMed ID: [31091251](#) doi:[10.1371/journal.pone.0216125](#)
- Fears, N.E., & Lockman, J.J. (2020). Using head-mounted eye-tracking to study hand-writing development. *Journal of Motor Learning and Development*, 8, 215–231. doi:[10.1123/jmld.2018-0057](#)
- Fiedler, K. (2017). What constitutes strong psychological science? The (neglected) role of diagnosticity and a priori theorizing. *Perspectives on Psychological Science*, 12(1), 46–61. PubMed ID: [28073328](#) doi:[10.1177/1745691616654458](#)
- Field, S.C., Esposito Bosma, C.B., & Temple, W.A. (2020). Comparability of the test of gross motor development– Second edition and the test of gross motor development– Third edition. *Journal of Motor Learning and Development*, 8, 107–125. doi:[10.1123/jmld.2018-0058](#)
- Flake, J.K., & Fried, E.I. (2019). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *PsyArXiv*. doi:[10.31234/osf.io/hs7wm](#)
- Forscher, B.K. (1963). Chaos in the brickyard. *Science*, 142(3590), 339. doi:[10.1126/science.142.3590.339](#)
- Freeman, J. (2018). LGBTQ scientists are still left out. *Nature*, 559, 27–28. PubMed ID: [29968839](#) doi:[10.1038/d41586-018-05587-y](#)
- Garcia-Masso, X., Estevan, I., Izquierdo-Herrera, R., Villarrasa-Sapiña, I., & Gonzalez, L.-M. (2020). Postural control profiles of typically developing children from 6 to 12 years old: A self-organizing maps approach. *Journal of Motor Learning and Development*, 8, 52–66. doi:[10.1123/jmld.2018-0016](#)
- Hagger, M.S., Chatzisarantis, N.L., Alberts, H., Anggono, C.O., Batailler, C., Birt, A.R., . . . Calvillo, D.P. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, 11(4), 546–573. doi:[10.1177/1745691616652873](#)
- Hauge, T.C., Katz, G.E., Davis, G.P., Jaquess, K.J., Reinhard, M.J., Costanzo, M.E., . . . Gentili, R.J. (2020). A novel application of Levenshtein distance for assessment of high-level motor planning underlying performance during learning of complex motor sequences. *Journal of Motor Learning and Development*, 8, 67–86. doi:[10.1123/jmld.2018-0060](#)
- Hedges, L.V. (1987). How hard is hard science, how soft is soft science? The empirical cumulativeness of research. *American Psychologist*, 42(5), 443–455. doi:[10.1037/0003-066X.42.5.443](#)
- Hooyman, A., Garbin, A., & Fisher, B. (2020). Paired associative stimulation rewired: A novel paradigm to modulate resting-state intracortical connectivity. *Journal of Motor Learning and Development*, 8, 174–193. doi:[10.1123/jmld.2018-0054](#)
- Immekus, J.C., Muntis, F., & Terson de Paleville, F. (2020). Predictor selection using lasso to examine the association of motor proficiency, postural control, visual efficiency, and

- behavior with the academic skills of elementary school-aged children. *Journal of Motor Learning and Development*, 8, 126–144. doi:[10.1123/jmld.2018-0023](https://doi.org/10.1123/jmld.2018-0023)
- John, L.K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23(5), 524–532. PubMed ID: [22508865](https://pubmed.ncbi.nlm.nih.gov/22508865/) doi:[10.1177/0956797611430953](https://doi.org/10.1177/0956797611430953)
- Kantak, S.S., & Winstein, C.J. (2012). Learning-performance distinction and memory processes for motor skills: A focused review and perspective. *Behavioural Brain Research*, 228(1), 219–231. PubMed ID: [22142953](https://pubmed.ncbi.nlm.nih.gov/22142953/) doi:[10.1016/j.bbr.2011.11.028](https://doi.org/10.1016/j.bbr.2011.11.028)
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9(6), 718–727. PubMed ID: [10607637](https://pubmed.ncbi.nlm.nih.gov/10607637/) doi:[10.1016/S0959-4388\(99\)00028-8](https://doi.org/10.1016/S0959-4388(99)00028-8)
- Kerr, N.L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, 2(3), 196–217. PubMed ID: [15647155](https://pubmed.ncbi.nlm.nih.gov/15647155/) doi:[10.1207/s15327957pspr0203_4](https://doi.org/10.1207/s15327957pspr0203_4)
- Lallensack, R. (2017). Female astronomers of colour face daunting discrimination. *Nature News*, 547(7663), 266–267. doi:[10.1038/nature.2017.22291](https://doi.org/10.1038/nature.2017.22291)
- Lataash, M.L. (2010). Motor synergies and the equilibrium-point hypothesis. *Motor Control*, 14(3), 294–322. PubMed ID: [20702893](https://pubmed.ncbi.nlm.nih.gov/20702893/) doi:[10.1123/mcj.14.3.294](https://doi.org/10.1123/mcj.14.3.294)
- Leek, J.T., & Peng, R.D. (2015). Statistics: P values are just the tip of the iceberg. *Nature News*, 520(7549), 612–612. doi:[10.1038/520612a](https://doi.org/10.1038/520612a)
- Lin, H., Saunders, B., Friese, M., Evans, N.J., & Inzlicht, M. (2019). Strong effort manipulations reduce response caution: A preregistered reinvention of the ego-depletion paradigm. *PsyArXiv*. doi:[10.31234/osf.io/dysa5](https://doi.org/10.31234/osf.io/dysa5)
- Lohse, K., Buchanan, T., & Miller, M. (2016). Underpowered and overworked: Problems with data analysis in motor learning studies. *Journal of Motor Learning and Development*, 4(1), 37–58. doi:[10.1123/jmld.2015-0010](https://doi.org/10.1123/jmld.2015-0010)
- Lohse, K.R., Shen, J., & Kozlowski, A.J. (2020). Modeling longitudinal outcomes: A contrast of two methods. *Journal of Motor Learning and Development*, 8, 145–165. doi:[10.1123/jmld.2019-0007](https://doi.org/10.1123/jmld.2019-0007)
- Lucca, K., Gire, D., Horton, R., & Sommerville, J.A. (2020). Automated measures of force and motion can improve our understanding of infants' motor persistence. *Journal of Motor Learning and Development*, 8, 194–214. doi:[10.1123/jmld.2019-0010](https://doi.org/10.1123/jmld.2019-0010)
- Makel, M.C., Plucker, J.A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, 7(6), 537–542. PubMed ID: [26168110](https://pubmed.ncbi.nlm.nih.gov/26168110/) doi:[10.1177/1745691612460688](https://doi.org/10.1177/1745691612460688)
- McShane, B.B., Gal, D., Gelman, A., Robert, C., & Tackett, J.L. (2019). Abandon statistical significance. *The American Statistician*, 73, 235–245. doi:[10.1080/00031305.2018.1527253](https://doi.org/10.1080/00031305.2018.1527253)
- Meehl, P.E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, 34(2), 103–115. doi:[10.1086/288135](https://doi.org/10.1086/288135)
- Mitchell, S., Loois, E.M., & Butterfield, S.A. (2020). A case for using hierarchical linear modeling in exercise science research. *Journal of Motor Learning and Development*, 8, 166–173. doi:[10.1123/jmld.2019-0003](https://doi.org/10.1123/jmld.2019-0003)
- Munafò, M.R., Nosek, B.A., Bishop, D.V., Button, K.S., Chambers, C.D., Du Sert, N.P., . . . Ioannidis, J.P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), 0021. doi:[10.1038/s41562-016-0021](https://doi.org/10.1038/s41562-016-0021)
- Newell, K.M. (1991). Motor skill acquisition. *Annual Review of Psychology*, 42(1), 213–237. doi:[10.1146/annurev.ps.42.020191.001241](https://doi.org/10.1146/annurev.ps.42.020191.001241)
- Ng, J.L., Button, C., Collins, D., Giblin, S., & Kennedy, G. (2020). Examining the factorial structure of the general movement competence assessment in children 8–10 years old. *Journal of Motor Learning and Development*, 8, 87–106. doi:[10.1123/jmld.2018-0047](https://doi.org/10.1123/jmld.2018-0047)

- Nosek, B.A., Beck, E.D., Campbell, L., Flake, J.K., Hardwicke, T.E., Mellor, D.T., . . . Vazire, S. (2019). Preregistration is hard, and worthwhile. *Trends in Cognitive Sciences*, 23(10), 815–818. PubMed ID: [31421987](#) doi:[10.1016/j.tics.2019.07.009](#)
- Nuzzo, R. (2015). How scientists fool themselves-and how they can stop. *Nature News*, 526(7572), 182–185. doi:[10.1038/526182a](#)
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. doi:[10.1126/science.aac4716](#).
- Patil, P., Peng, R.D., & Leek, J.T. (2016). What should researchers expect when they replicate studies? A statistical view of replicability in psychological science. *Perspectives on Psychological Science*, 11(4), 539–544. PubMed ID: [27474140](#) doi:[10.1177/1745691616646366](#)
- Poggio, T. (2012). The levels of understanding framework, revised. *Perception*, 41(9), 1017–1023. PubMed ID: [23409366](#) doi:[10.1068/p7299](#)
- Popper, K. (2005). *The logic of scientific discovery*. New York, NY: Routledge.
- Prinz, F., Schlange, T., & Asadullah, K. (2011). Believe it or not: How much can we rely on published data on potential drug targets? *Nature Reviews Drug Discovery*, 10(9), 712–712. PubMed ID: [21892149](#) doi:[10.1038/nrd3439-c1](#)
- Schmidt, R.A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82(4), 225–260. doi:[10.1037/h0076770](#)
- Simmons, J.P., Nelson, L.D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. PubMed ID: [22006061](#) doi:[10.1177/0956797611417632](#)
- Singal, A.G., Higgins, P.D., & Waljee, A.K. (2014). A primer on effectiveness and efficacy trials. *Clinical and Translational Gastroenterology*, 5(1), e45. doi:[10.1038/ctg.2013.13](#)
- Snijders, T.A.B., & Bosker, R.J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Thousand Oaks, CA: Sage Publishers.
- Szollosi, A., Kellen, D., Navarro, D., Shiffrin, R., van Rooij, I., Van Zandt, T., & Donkin, C. (2019). Preregistration is redundant, at best. *PsyArXiv*. doi:[10.31234/osf.io/x36pz](#)
- Thelen, E., & Smith, L.B. (1996). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT press.
- Vazire, S. (2017). Quality uncertainty erodes trust in science. *Collabra: Psychology*, 3(1), 1. doi:[10.1525/collabra.74](#)
- Wang, P., Infurna, F.J., & Schaefer, S.Y. (2020). Predicting motor skill learning in older adults using visuospatial performance. *Journal of Motor Learning and Development*, 8, 38–51. doi:[10.1123/jmld.2018-0017](#)
- Wasserstein, R.L., & Lazar, N.A. (2016). The ASA’s statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2), 129–133. doi:[10.1080/00031305.2016.1154108](#)
- Wasserstein, R.L., Schirm, A.L., & Lazar, N.A. (2019). Moving to a world beyond “ $p < .05$ ”. *The American Statistician*, 73(51), 1–19. doi:[10.1080/00031305.2019.1583913](#)
- Wigboldus, D.H., & Dotsch, R. (2016). Encourage playing with data and discourage questionable reporting practices. *Psychometrika*, 81(1), 27–32. PubMed ID: [25820979](#) doi:[10.1007/s11336-015-9445-1](#)
- Woolf, S.H. (2008). The meaning of translational research and why it matters. *Journal of the American Medical Association*, 299(2), 211–213. PubMed ID: [18182604](#)
- World Health Organization. (2001). *International classification of functioning, disability and health: ICF*. Geneva, Switzerland: Author.
- Wulf, G., & Lewthwaite, R. (2016). Optimizing performance through intrinsic motivation and attention for learning: The OPTIMAL theory of motor learning. *Psychonomic*

Bulletin & Review, 23(5), 1382–1414. PubMed ID: [26833314](#) doi:[10.3758/s13423-015-0999-9](#)

Yoder, J.B., & Mattheis, A. (2016). Queer in STEM: Workplace experiences reported in a national survey of LGBTQA individuals in science, technology, engineering, and mathematics careers. *Journal of Homosexuality*, 63(1), 1–27. PubMed ID: [26241115](#) doi:[10.1080/00918369.2015.1078632](#)