

Week 2 — Theory & conceptual clarity

How well is the field defining its constructs, using theory generatively rather than rhetorically, and avoiding the conflation of distinct concepts?

Objective evidence review · companion to Week 2 · prepared July 2026

Scope & method. Rapid narrative review, not a systematic review. Structured search of the sport-science, sport-psychology and measurement literature (PubMed, Google Scholar), together with two Sports Science Replication Centre (SSRC) resources supplied for this review: a computational audit of theory use (Warne, Gorman et al., in preparation) and the FAIR Theory Database (an internal catalogue). Unpublished and internal sources are flagged as such; peer-reviewed figures are reported as the cited studies state them.

The principle under review

WEEK 2 · CONCEPTUAL CLARITY AND GENERATIVE THEORY

Week 2 argues that a claim must be defined precisely enough to be tested, that abstract constructs must be distinguished from the measures that index them, and that a theory earns its place by generating specific, falsifiable predictions rather than by lending rhetorical legitimacy. The empirical questions are whether the field uses theory generatively, defines its constructs consistently, and clarifies terms when they become ambiguous.

1 · How well is it currently applied?

THE EVIDENCE

9% of theories identified in 269 sports-science articles were <i>operational</i> (generated a testable prediction) — Warne, Gorman et al. (in prep.)	1% of articles showed strong hypothesis–theory alignment (all hypotheses derived from theory) — same audit	5% of articles fully re-engaged with their theoretical framework in the discussion — same audit
--	--	---

A direct audit finds theory use is rhetorical, not generative

The most direct evidence comes from a large-scale computational audit conducted by the SSRC (Warne, Gorman et al., in preparation; unpublished). Using an automated, deterministic large-language-model pipeline to analyse the introductions and discussions of 269 applied sports-science articles, the audit identified 421 theoretical-framework instances (mean 1.6 per article; 93% explicitly named). Only 9% of these theories were classified as *operational* — directly generating the specific prediction the study tested. Strong hypothesis–theory alignment, in which all hypotheses were derived from the stated theory rather than merely stated as consistent with it, occurred in just 1% of articles, and full re-engagement with the theoretical framework in the discussion in only 5%. The authors conclude that theory use in the field is “more rhetorical than generative.” (As this is an in-preparation working paper, figures should be treated as provisional pending peer review.)

An internal catalogue shows uneven theoretical development

The SSRC's FAIR Theory Database — an internal, machine-readable catalogue of named theories in sport and exercise science — scores each theory on operationalisation clarity, replication support, cumulative development, precision/falsifiability, and domain breadth (a Lakatosian and Meehlian framework). Across its 97 catalogued theories, 36 reach the highest tier while 48 sit in the middle tier and 13 in the lowest — that is, roughly 63% show gaps in operationalisation, replication, or falsifiability that place them below the top tier. This is an internal resource rather than a peer-reviewed dataset, but it is consistent with the audit above: many named theories in the field are not yet developed to the point of generating precise, testable predictions.

Theory use is inconsistent even where quantified in adjacent work

The broader open-science literature reaches the same conclusion from a different angle. Van Lissa and colleagues (2026) argue that a decade of reform has improved the “back end” of research (data, code, analysis) while leaving theory specification — the front end — informal, verbal and non-interoperable, so that the same theoretical label routinely covers different operationalisations across studies. In health-

behaviour research, Prestwich and colleagues (2014) found that only 56% of 190 physical-activity and dietary interventions reported any theory base, and that where theory was named it was often applied loosely rather than used to select intervention components.

Construct definitions proliferate without consensus

Several prominent constructs lack agreed definitions or measures. After roughly three decades of research, mental toughness still shows persistent disagreement over whether it is uni- or multi-dimensional, a trait or a state, and how it should be operationalised (Gucciardi, Hanton, Gordon, Mallett & Temby, 2015). A clear jangle fallacy comes from the adjacent achievement literature: Credé, Tynan and Harms (2017), meta-analysing 88 samples (66,807 individuals), found “grit” correlates with conscientiousness at $r = 0.84$ — strong enough to question whether it is distinct at all.

Conceptual clarification is achievable — and happening

The field also contains clear positive examples. “Fatigue,” long used as one word for distinct phenomena, was decomposed by Enoka and Duchateau (2016) into performance and perceived fatigability. “Training load” has been repeatedly clarified by Impellizzeri, Marcora and Coutts (2019) — external versus internal load — and extended by Impellizzeri and colleagues (2023) using exposure–dose concepts from epidemiology. Distinguishing constructs from their measures, however, is still frequently neglected: Flake and Fried (2020) document widespread questionable measurement practices, and Impellizzeri and Marcora (2009) note that surrogate and field measures are often adopted without validation (developed in Week 5).

Source	Finding	Implication
Warne, Gorman et al. (in prep.) <i>SSRC working paper</i> — unpublished	Audit of 269 articles: 9% of theories operational; 1% strong hypothesis–theory alignment; 5% discussion re-engagement	Theory use is rhetorical rather than generative
FAIR Theory Database <i>SSRC internal catalogue</i>	97 theories scored; ~63% below the top development tier (48 mid, 13 low)	Theoretical development in the field is uneven
Van Lissa et al. (2026) <i>Persp. Psych. Science (adjacent)</i>	Theory specification remains informal and non-interoperable despite open-science reform	Same label covers divergent operationalisations
Prestwich et al. (2014) <i>Health Psychology</i>	Only 56% of 190 PA/diet interventions reported a theory base; theory applied loosely	Theory use is inconsistent
Gucciardi et al. (2015) <i>J. Personality</i>	No consensus on dimensionality, traitness or measurement of mental toughness	A popular construct remains poorly defined
Credé, Tynan & Harms (2017) <i>JPSP (adjacent)</i>	“Grit” correlates with conscientiousness at $r = 0.84$ (88 samples)	Jangle fallacy: a “new” construct duplicates an old one
Enoka & Duchateau (2016) <i>MSSE</i>	Decomposed “fatigue” into performance vs perceived fatigability	Positive example: resolving a jingle fallacy
Impellizzeri, Marcora & Coutts (2019); Impellizzeri et al. (2023) <i>IJSP; Sports Medicine</i>	Re-clarified and refined external vs internal training load	Positive example: active conceptual clarification
Flake & Fried (2020); Impellizzeri & Marcora (2009) <i>AMPPS; IJSP</i>	Measures used without validity evidence; scores treated as constructs	Construct–measure conflation is common

2 · Is there cause for concern?

OBJECTIVE ASSESSMENT

ASSESSMENT **Substantial concern**

The direct audit is the decisive evidence: if only about 9% of cited theories generate a specific tested prediction and only 1% of articles show full hypothesis–theory alignment, then the field is largely not using theory in the generative sense the week describes. This is reinforced by an internal catalogue showing uneven theoretical development, by adjacent evidence that theory specification remains informal, and by persistent definitional problems for popular constructs. Because conceptual and theoretical clarity is upstream of measurement, design and analysis, these weaknesses propagate into the replication problems documented in Week 1.

Two qualifications keep the assessment proportionate. First, the strongest single figure comes from an unpublished, in-preparation audit and should be treated as provisional until peer-reviewed, though it is consistent with independent lines of evidence. Second, the field contains clear, successful examples of conceptual clarification (fatigue, training load), showing the problem is tractable and requires rigorous definition rather than additional resources.

3 • Immediate steps to improve

ACTIONABLE

1. **Use theory operationally, not rhetorically:** derive the specific hypothesis from the named theory's core claims, rather than citing the theory only for background.
2. **Define the construct explicitly, in one sentence,** and state the bridge assumption linking each measure to it, with validity evidence (Week 5).
3. **Re-engage with the theory in the discussion:** interpret results in terms of what they mean for the theory, including how a null result would bear on it.
4. **Check for jingle and jangle** before adopting a construct, and use established, defined terminology rather than umbrella terms.
5. **Specify boundary conditions and mechanisms** so the prediction is constraining (not merely a directional association that any theory could accommodate).

References

- Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: a meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*, 113(3), 492–511. <https://doi.org/10.1037/pspp0000102>
- Enoka, R. M., & Duchateau, J. (2016). Translating fatigue to human performance. *Medicine & Science in Sports & Exercise*, 48(11), 2228–2238. <https://doi.org/10.1249/MSS.0000000000000929>
- FAIR Theory Database (2026). Internal machine-readable catalogue of theories in sport and exercise science (v1.3, 97 theories). Sports Science Replication Centre / OpenClaw Research Ecosystem. *Internal resource*.
- 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>
- Gucciardi, D. F., Hanton, S., Gordon, S., Mallett, C. J., & Temby, P. (2015). The concept of mental toughness: tests of dimensionality, nomological network, and traitness. *Journal of Personality*, 83(1), 26–44. <https://doi.org/10.1111/jopy.12079>
- Impellizzeri, F. M., & Marcora, S. M. (2009). Test validation in sport physiology: lessons learned from clinimetrics. *International Journal of Sports Physiology and Performance*, 4(2), 269–277. <https://doi.org/10.1123/ijssp.4.2.269>
- Impellizzeri, F. M., Marcora, S. M., & Coutts, A. J. (2019). Internal and external training load: 15 years on. *International Journal of Sports Physiology and Performance*, 14(2), 270–273. <https://doi.org/10.1123/ijssp.2018-0935>
- Impellizzeri, F. M., Shrier, I., McLaren, S. J., et al. (2023). Understanding training load as exposure and dose. *Sports Medicine*, 53(9), 1667–1679. <https://doi.org/10.1007/s40279-023-01833-0>
- Prestwich, A., Sniehotta, F. F., Whittington, C., Dombrowski, S. U., Rogers, L., & Michie, S. (2014). Does theory influence the effectiveness of health behavior interventions? Meta-analysis. *Health Psychology*, 33(5), 465–474. <https://doi.org/10.1037/a0032853>
- Van Lissa, C. J., Peikert, A., Ernst, M. S., van Dongen, N. N. N., Schönbrodt, F. D., & Brandmaier, A. M. (2026). To be FAIR: theory specification needs an update. *Perspectives on Psychological Science*. <https://doi.org/10.1177/17456916251401850>
- Warne, J., Gorman, B., et al. (in preparation). Theory use in applied sports science: a computational audit of explicit and implicit theoretical frameworks, hypothesis derivation, and discussion re-engagement. *Sports Science Replication Centre working paper, Technological University Dublin. Unpublished*.

Improving Research Practices in Sport & Exercise Science — a free short course for the Sports Science Replication Centre. Current State Review, Week 2. Unpublished and internal SSRC sources flagged; peer-reviewed figures reported as stated in the cited sources; adjacent-field examples flagged. For education.