FAIR Theory: Applying Open Science Principles to the Construction and Iterative Improvement of Scientific Theories

Caspar J. Van Lissa[1], Aaron Peikert[2,3], Andreas M. Brandmaier[2,3,4], & Felix Schönbrodt[5]

[1] Tilburg University, dept. Methodology & Statistics

[2] Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany

[3] Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany

[4] Department of Psychology, MSB Medical School Berlin, Berlin, Germany

[5] Other affiliations

Author Note

This is a preprint paper, generated from Git Commit # 4d935da1.

The authors made the following contributions. Caspar J. Van Lissa: Conceptualization, Formal Analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Writing – original draft, Writing – review & editing; Aaron Peikert: Formal Analysis, Writing – original draft, Writing – review & editing; Andreas M. Brandmaier: Formal Analysis, Writing – original draft, Writing – review & editing; Felix Schönbrodt: Conceptualization, Writing – review & editing.

Correspondence concerning this article should be addressed to Caspar J. Van Lissa, Professor Cobbenhagenlaan 125, 5037 DB Tilburg, The Netherlands. E-mail: c.j.vanlissa@tilburguniversity.edu

Abstract

Test test.

Keywords: meta theory, theory formation, cumulative science, formal models

Word count: 3547

FAIR Theory: Applying Open Science Principles to the Construction and Iterative Improvement of Scientific Theories

The FAIR Guiding Principles (hereafter: FAIR principles) were established to make research data more Findable, Accessible, Interoperable and Reusable [REF]. Since their inception, scholars have demonstrated their relevance for making other digital research artefacts more open. This paper argues that the FAIR principles can advance effective and transparent scholarly communication about theory as well. To this end, we introduce "FAIR theory": a digital representation of theory compliant with the FAIR principles. FAIR Theory has the potential to substantially advance the efficiency of scholarly communication and accelerate cumulative knowledge acquisition.

The Need for FAIR Theory

The so-called "replication crisis" has prompted extensive reforms in social science (Lavelle, 2021; Scheel, 2022). Concern that undisclosed flexibility in analyses was to blame for the abundance of non-replicable findings led to widespread adoption of open science practices like preregistration and replication (Nosek et al., 2015). These various practices ensure transparent and repeated testing of hypotheses. However, recent reviews show that most preregistered hypothesis tests are not supported (Scheel, Schijen, & Lakens, 2021). Increased rigor in testing has revealed that the root cause of the replication crisis is more fundamental: psychological theories rarely produce hypotheses that are found to be true, and are often sufficiently vague to explain contradictory findings.

Scholars have raised concerns about the state of theory in social science for nearly 50 years (Meehl, 1978; Robinaugh, Haslbeck, Ryan, Fried, & Waldorp, 2021). One concern is that social scientific theories often lack the precision and clarity of theories in the physical sciences, otherwise known as *formalization* (Szollosi & Donkin, 2021). A second concern, which received less attention, is the lack of transparent and democratic scholarly communication about psychological theory. The present paper focuses on the second concern, advancing transparent communication about theory by applying open science principles to psychological theory for the first time and introducing the concept of *FAIR Theory*.

Theory and Scientific Progress

According to the *empirical cycle* (de Groot, 1961), a philosophical model of cumulative knowledge acquisition, research ideally follows a cyclical process with two phases (Figure 1). In the deductive phase, hypotheses derived from theory are tested on data. In the inductive phase, patterns observed in data are generalized to theoretical principles. In this model, theories are the vehicle of scientists' understanding of phenomena. Ideally, they are iteratively updated based on deductive testing and inductive theory construction.

In a progressive research program (Lakatos, 1971), the cycle is regularly completed to iteratively advance our understanding of the studied phenomena. There are indications that contemporary psychology falls short of this idealized model, however. Firstly, because deductive research is over-represented in the literature. According to one estimate, 89.6% of published studies tests hypotheses, which suggests that the literature is predominantly comprised of deductive research (Kühberger, Fritz, & Scherndl, 2014). Closer examination reveals, however, that the link between theory and hypothesis is often tenuous (Oberauer & Lewandowsky, 2019; Scheel, Tiokhin, Isager, & Lakens, 2021). Only 15% of deductive studies reference theory at all (McPhetres et al., 2021). This invites the question where the other hypotheses come from, and to what extent testing an ad-hoc hypothesis not grounded in theory can advance our principled understanding of the studied phenomena. These statistics suggest that theory has an uncomfortable and paradoxical role in contemporary psychology: The majority of papers ostensibly test hypotheses, but these are rarely derived from theory, and test results rarely contribute to the improvement of existing theories. Consequently, theories either persist unchanged, immune to refutation, or are forgotten [REF Meehl].

The Solution

Applying the FAIR principles to scientific theories facilitates transparent scholarly communication and accelerates cumulative knowledge acquisition for several reasons.

Findability. One major obstacle to cumulative knowledge acquisition is the difficulty in finding relevant theories. The primary unit of dissemination and search in psychology is often the academic paper. A paper may contain multiple resources - including materials, data, code, and theory - but there is no unified search engine for theory, or standardized metadata to signal the presence of theory within a work, nor even a standardized keyword (model, framework, etc are often used interchangeably with theory). Making theories more findable would allow them to be systematically indexed, tagged, and made searchable. This would allow researchers to easily identify relevant theories to inform their hypotheses, grounding their work in established theoretical foundations. Making theories findable also increases the impact and reuse potential of theories across scientific contexts. For example, it becomes possible to study theories (the same way libraries and search engines have enabled scholars to study the literature via systematic reviews); theories might have transdisciplinary impact; and the structure of theories from one discipline might inspire analogical modeling in other fields, the way - for example, predator-prey models have inspired theories of XXX and the Eysenck model of atomic magnetism has inspired a network theory of depression.

One effort that improves theories' findability is post-hoc curation. For example, Gray and colleagues introduced a format for representing theories, and post many examples on their website [REF GRAY]. Similarly, Borsboom and colleagues seek to establish a dictionary of psychological "phenomena" (which are not strictly theories, but are patterns reliably evidenced by data that theory should seek to explain) [REF BORSBOOM]. Another method that could ameliorate theory Findability is modular publishing. In this approach, distinct resources (data, code, instruments, theory) are individually published as citable academic assets. Such output can be linked to a specific paper.

Accessibility. Advancing transparent scholarly communication about theory requires that theories are accessible to all researchers and other stakeholders. Publishing behind paywalls clearly creates obstacles. Open Access publishing increases the accessibility of all academic output, including theory. If theories were more accessible, researchers could reuse, refine, and relate their work to those theories, thus accelerating cumulative knowledge acquisition. Making theories more accessible also allows other stakeholders (e.g., practitioners, policy makers, advocates) to inform themselves of the current scientific understanding of specific phenomena: More so than single studies, theories offer a top-down, big picture representation of the phenomena studied in a field.

Beyond open access publishing, making theories accessible also requires communicating them clearly and explicitly. This property has been dubbed "discursive survival [...], the ability to be understood" (**guestWhatMakesGood2024?**). Yet contemporary psychological theory is often ambiguous [REF Frankenhuis]. Ambiguity introduces a dependency on the original author for clarification, because the author can always claim that his ideas were misconstrued or misrepresented. This discourse on "Great Man Theorizing" touches upon the problems this introduces [REF Guest et al moral theory]. For example, dependency on interpretation by the author creates a potential for gatekeeping - the author could insist that work requires their involvement, which violates checks and balances of scientific research. Moreover, if a theory is refuted, its author could claim that the author of the refuting paper did not interpret the theory correctly. To some extent, this relates to the problem of translation [REF Duhem]: it is not possible to entirely formalize an idea to enable unambiguous interpretation. Nonetheless, taking care to formalize a theory to the maximum extent possible advances Accessibility.

There are clear indications that psychological theory has limited interoperability. For one, theories are rarely *"refuted nor corroborated, but instead merely fade away as people lose interest"* (Meehl, 1978). To be interoperable, psychological theory would need to play a concrete part in scientists' day-to-day work. For example, it should be possible to integrate theory directly into analysis workflows: to derive hypotheses, select control variables, and guide model specification. The aforementioned lack of formalization [REF Robinaugh] and ambiguity [REF Frankenhuis] prevent interoperability in such a practical sense. Additionally, theories should be interoperable with each other: for instance, it should be possible to embed a specific theory about the process of emotion regulation [REF Gross] within a theory of emotion regulation development [REF Morris].

The ultimate goal of the FAIR principles is to facilitate the reuse digital research artefacts. While theory is ostensibly held in high regard in the social sciences, there also appears to be a norm *against* theory reuse. This is evident from the quip that *"[Theories are] like toothbrushes — no self-respecting person wants to use anyone else's"* (Mischel, 2008). In our workshops on FAIR Theory, we often hear questions such as "who owns theory", and "who determines how a theory may be reused or changed"? Licensing theories for reuse provides an unambiguous answer to such questions.

One important consideration is that there may be legal limits to the rights authors can claim over theory. An important concept in copyright law is the *idea-expression dichotomy* [REF].

o. Ideas, facts, and concepts are not protected by Although they are not protectable by copyright, the expression of those ideas, facts, and concepts are protectable, such as in a description, explanation, or illustration or as a database of facts. Separating an idea from the expression or manifestation of that idea is known in copyright law as the Copyright explicitly does not *"extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery"*, but it may extend to creative works expressing that idea (e.g., writing, visual illustrations). Copyright pertaining to recipes offer an interesting perspective: Whereas the recipes themselves are not protected, *"substantial literary expression in the form of an explanation or directions"* can be protected by copyright. Extrapolating to FAIR Theory, it would seem that vague, ambiguous verbal explanations of theories are more likely to qualify for copyright protection than formal theories. If copyright does not cover (formal) theory, then it would be pointless to adopt a license that assumes copyright protection. Furtermore, even if copyright would apply, research is covered under “fair use” excemptions to copyright law, so copyright would impose few restrictions to theory reuse in scholarly communication. In sum, the CC0 (no rights reserved) license seems most appropriate for FAIR Theory; it explicitly waives all rights and encourages reuse. Simultaneously, to comply with the interoperability principle, the principle of comprehensive citation [REF TOP guidelines], and the spirit of good scholarship, it is essential that scholars do comprehensively cite theory, including prior work that new theories are based on.

3. Interoperability: Facilitating Integration Across Fields In many scientific disciplines, researchers often operate in isolated silos, using different terminologies, methodologies, and conceptual frameworks. The interoperability aspect of FAIR focuses on the ability of data (or in this case, theories) to be integrated with other systems using common standards, formats, and vocabularies.

How it helps: If theories are made interoperable, they can be used in multi-disciplinary research and can better interface with diverse fields. This reduces duplication of effort and enables researchers from different areas to combine and compare theories, bridging gaps between disciplines and fostering innovative, cross-disciplinary approaches. Researchers can easily synthesize theoretical perspectives, generating new hypotheses that are more comprehensive and grounded in a wider range of evidence. 4. Reusability: Enabling Continuous Improvement of Theories For knowledge to accumulate meaningfully, theories need to be adaptable and subject to continuous revision as new data emerge. The reusability principle calls for thorough documentation and community standards to ensure that data (or theories) can be applied in new contexts.

How it helps: Reusable theories would provide clear metadata, version histories, and supporting evidence, allowing other researchers to test, refine, or extend them as new empirical findings emerge. This improves the feedback loop between theory and experimentation, ensuring that test results inform future theoretical development. It encourages the use of theories not just for their original purpose but in new and evolving contexts, thus ensuring that the theory remains dynamic and continuously refined over time. Conclusion: Making Theories FAIR Accelerates Scientific Progress Adopting the FAIR principles for theories can address key challenges in the current research landscape, where theories often remain isolated and underutilized. By making theories findable, accessible, interoperable, and reusable, researchers can ensure that their work is grounded in a shared, transparent, and cumulative body of knowledge. This approach enhances scholarly communication, allowing for greater scrutiny, replication, and collaboration across disciplines, ultimately leading to faster, more reliable, and more impactful scientific progress.

Publication is not Enough

Merely publishing a theory does not make it open; to be open, theory should adhere to established open science standards. The FAIR principles, initially introduced as a standard for open research data, have since been applied to other forms of digital scholarly output (e.g., software Lamprecht et al., 2019). We propose to apply the FAIR principles to digital representations of theory as well, introducing a FAIR metadata format to represent (formal) theories. The resulting theories are made *Findable* via a DOI, Accessible in a machine- and human-readable filetype, Interoperable within the data analysis environment, and Reusable in the practical and legal sense, so that they may be improved over time.

Adapting the FAIR Principles

The FAIR Principles were devised to make scholarly data more findable, accessible, interoperable, and reusable. From their inception, these principles were developed with “other research resources” in mind. Scholars have translated the FAIR principles to, e.g., research software [REF Lamprecht]. The present paper further extends the FAIR principles' definition to theory, see Table 1.

FAIR Theory and Recognition & Rewards

In the spirit of DORA, extending the FAIR principles to theory helps researchers obtain credit for their theoretical contributions - obviating the necessity of publishing a theoretical paper, which can be challenging. From a meta-science perspective, FAIR theory facilitates studying the state of theory in a particular subfield, and comparing theories' substantive and structural properties. Version control and cross-referencing additionally enable tracing and studying the ancestry and development of theories.

FAIR theory provides a clear deliverable, and a clear goal, for scholars and institutions seeking to promote contributions to theory.

There are key distinctions between theory and other FAIR digital research artefacts. With this in mind, following the example of Lamprecht and colleagues, we reflect on how the criteria underlying the FAIR Principles apply to theory.

The Role of Theory Formalization

Concerns about the state of theory are a recurring theme in the psychological literature, but previous writing has focused on theory formalization as a solution for ambiguity in psychological theory. Greater formality increases theories' *empirical content*, making them easier to falsify, which necessitates revising them, thus advancing our principled understanding of the phenomena they describe. Conceptually, theory formalization is orthogonal to FAIR theory. FAIR Theory does not require theories to be formal, and formal theory can be represented in a way that is not FAIR. It is - in principle - possible to represent a collection of verbal statements as a FAIR Theory. While FAIR Theory is fully consistent with formal theory, it does not require theories to be formal.

Version Control

- One source of potential improvements of theory methodology that has not been previously considered is computer science.
- The process of “iteratively improving” digital objects - in this case, computer code - is well understood.
- Recent work like the FAIR software principles has demonstrated that ideals of open science apply to computer science as well.
- This paper argues that, conversely, principles of computer science - particularly version control, algorithmic hypothesis generation (find better word; this is about using the digital theory object to derive implied hypotheses), and integrated testing, can also be used to improve theory methods in the social science.
- We introduce “FAIR theory”, a digital research artifact to represent formal social scientific theories
- FAIR theory can be version controlled; any time new insights require modifications of the theory, these modifications can be documented in a traceable and reversable manner. Version control also enables diffuse collaboration in theory development, as other researchers can submit “pull requests” to suggest modifications of a theory, or can “fork” existing theories to create a spin-off from an existing theory.

Examples

Formalizing the Empirical Cycle

In this example, we represent the empirical cycle - a theory of cumulative knowledge production through scientific research - as FAIR theory. As several authors have taken inspiration from the work by De Groot, we compare our interpretation of the original theory to the interpretation of others. Originally, the theory has the following structure:

```
digraph {

  observation;
  induction;
  deduction;
  test;
  evaluation;

  observation -> induction;
  induction -> deduction;
  deduction -> test;
  test -> evaluation;
  evaluation -> observation;

}
```

Subsequently, Wagenmakers and colleagues modified the theory by *"[adding the] Whewell-Peirce-Reichenbach distinction between the context of discovery and the context of justification"*:

```
digraph {

  subgraph cluster_discovery {
    label="Discovery";
    hypothesis [label="New hypothesis"];
    prediction [label="New prediction"];
  }
  data  [label="Old knowledge and old data"];
```

```
  subgraph cluster_justification {
    label="Justification";
    test [label="Test on new data"];
    evaluation;
  }

  data -> hypothesis [label="Speculate & explore"];
  hypothesis -> prediction  [label="Deduce"];
  prediction -> test  [label="Design new experiment"];
  test -> evaluation  [label="Statistical analysis"];
  evaluation -> data  [label="Knowledge accumulation"];

}
```

Note, however, that there appear to be further changes: the phases of the cycle have been renamed, and the annotations suggest a move towards experimental empirical psychology that was absent in the original formulation. Moreover, the label "knowledge accumulation" invites the question of exactly *how* knowledge accumulates upon evaluation of a prior experiment. As this lack of cumulative knowledge acquisition appears to be precisely where contemporary research practice falls short, this ambiguity invites further improvement of the theory.

Our work, too is inspired by De Groot, but our take on the empirical cycle is different again:

```
digraph {

  theory;
  prediction;
```

```
  test [label="inferential procedure"];
  observation;

  theory -> prediction [label="deduction"];
  prediction -> test;
  test -> observation;
  observation -> theory [label="generalization"];

}
```

In our representation, induction is not a separate phase but a mode of reasoning by which specific observations are generalized into theory. For example, the refutation of a hypothesized effect, or the serendipitous observation of some pattern in data, might be a reason to revise or construct theory. Induction, incidentally, also occurs within the link from prediction to testing: in the form of the inductive bias of methods used to perform the test, and auxiliary assumptions that must be made to address remaining theoretical ambiguities.

Using FAIR Theory to Perform Causal Inference

Some have argued that *causal explanations* are a property of good theory [REF Meehl, etc?]. According to Pearl and colleagues, explicit assumptions about the direction of causality allow one to perform causal inference even on cross-sectional data. Any formal theory that is explicit about direction of causality could thus be used to guide causal inference, and could even be integrated into the analysis environment.

In this example, we illustrate how to use DAGs for causal inference, including the detection of a violation of the initial model and subsequent adaptation of the DAG. We could use that to illustrate updating FAIR theory:

https://currentprotocols.onlinelibrary.wiley.com/doi/full/10.1002/cpz1.45

We can find more examples of causal inference with DAGs in these tutorials:

https://www.r-bloggers.com/2019/08/causal-inference-with-dags-in-r/

https://www.r-bloggers.com/2018/08/applications-of-dags-in-causal-inference/

- Theory is the vehicle of cumulative knowledge acquisition
- According to the empirical cycle, ideally, hypotheses are derived from theory, then tested in data, and theory is amended based on the resulting insights. When this cycle is regularly completed, theories become ever more veracious representations of social scientific phenomena.
- At present, there is concern over a theory crisis in the social sciences, which highlights that this system is not functioning as intended, and highlights the need for better theory.
- One source of potential improvements of theory methodology that has not been previously considered is computer science.
- The process of "iteratively improving" digital objects - in this case, computer code - is well understood.
- Recent work like the FAIR software principles has demonstrated that ideals of open science apply to computer science as well.
- This paper argues that, conversely, principles of computer science - particularly version control, algorithmic hypothesis generation (find better word; this is about using the digital theory object to derive implied hypotheses), and integrated testing, can also be used to improve theory methods in the social science.
- We introduce "FAIR theory", a digital research artifact to represent formal social scientific theories
- FAIR theory can be version controlled; any time new insights require modifications of the theory, these modifications can be documented in a traceable and reversable manner. Version control also enables diffuse collaboration in theory development, as

other researchers can submit “pull requests” to suggest modifications of a theory, or can “fork” existing theories to create a spin-off from an existing theory.

- FAIR theory allows for algorithmic derivation of hypotheses implied by the theory.
- FAIR theory enables integration testing: researchers can build a “test suite” of evidence that must be explainable by the theory, and any modifications of the theory must also pass the test suite.
- To illustrate FAIR theory’s potential to accelerate cumulative knowledge acquisition, we present several tutorial examples, developed in collaboration with applied researchers across fields of social science.

Discussion

Future Directions

One remaining issue that intersects with FAIR Theory is the measurement and operationalization of psychological constructs. Aside from the aforementioned “theory crisis”, there has been talk of a “measurement crisis”: it is not always clear how theoretical constructs are operationalized, and many existing instruments have poor psychometric properties [REF]. Additionally, the “jingle-jangle” fallacy is prevalent in the social sciences: the same term is often used for distinct constructs, and conversely, different terms are used to refer to the same construct. FAIR Theory can help address the measurement crisis: since theories can reference other theories and resources, it is possible to extend a structural theory with a theory of

FAIR Theory incorporates theory into open science workflows, facilitates scholarly communication about theories, making it easier to share theories with less opportunity for ambiguity and misunderstanding. FAIR Theories are easier to find, and facilitate sharing, reusing, and updating open theories. More efficient and transparent communication about theory democratizes and accelerates cumulative knowledge acquisition, removes barriers for knowledge exchange with the global scholarly community, opens theory development to diverse perspectives, and enables (distributed and adversarial) collaboration.

References

de Groot, A. D. (1961). *Methodologie: Grondslagen van onderzoek en denken in de gedragswetenschappen.* 's Gravenhage: Uitgeverij Mouton. Retrieved from https://books.google.com?id=6hiBDwAAQBAJ

Kühberger, A., Fritz, A., & Scherndl, T. (2014). Publication Bias in Psychology: A Diagnosis Based on the Correlation between Effect Size and Sample Size. *PLoS ONE*, *9*(9), e105825. https://doi.org/10.1371/journal.pone.0105825

Lakatos, I. (1971). History of Science and its Rational Reconstructions. In R. C. Buck & R. S. Cohen (Eds.), *PSA 1970: In Memory of Rudolf Carnap Proceedings of the 1970 Biennial Meeting Philosophy of Science Association* (pp. 91–136). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-3142-4_7

Lamprecht, A.-L., Garcia, L., Kuzak, M., Martinez, C., Arcila, R., Martin Del Pico, E., … Capella-Gutierrez, S. (2019). Towards FAIR principles for research software. *Data Science*, 1–23. https://doi.org/10.3233/DS-190026

Lavelle, J. S. (2021). When a Crisis Becomes an Opportunity: The Role of Replications in Making Better Theories. *The British Journal for the Philosophy of Science*, 714812. https://doi.org/10.1086/714812

McPhetres, J., Albayrak-Aydemir, N., Mendes, A. B., Chow, E. C., Gonzalez-Marquez, P., Loukras, E., … Volodko, K. (2021). A decade of theory as reflected in Psychological Science (2009–2019). *PLOS ONE*, *16*(3), e0247986. https://doi.org/10.1371/journal.pone.0247986

Meehl, P. E. (1978). Theoretical Risks and Tabular Asterisks: Sir Karl, Sir Ronald, and the Slow Progress of Soft Psychology. *Journal of Consulting & Clinical Psychology*, *46*(4), 806–834.

Mischel, W. (2008). The Toothbrush Problem. *APS Observer*, *21*. Retrieved from https://www.psychologicalscience.org/observer/the-toothbrush-problem

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., …

Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, *26*(5), 1596–1618. https://doi.org/10.3758/s13423-019-01645-2

Robinaugh, D. J., Haslbeck, J. M. B., Ryan, O., Fried, E. I., & Waldorp, L. J. (2021). Invisible Hands and Fine Calipers: A Call to Use Formal Theory as a Toolkit for Theory Construction. *Perspectives on Psychological Science*, *16*(4), 725–743. https://doi.org/10.1177/1745691620974697

Scheel, A. M. (2022). Why most psychological research findings are not even wrong. *Infant and Child Development*, *31*(1), e2295. https://doi.org/10.1002/icd.2295

Scheel, A. M., Schijen, M. R. M. J., & Lakens, D. (2021). An Excess of Positive Results: Comparing the Standard Psychology Literature With Registered Reports. *Advances in Methods and Practices in Psychological Science*, *4*(2), 25152459211007467. https://doi.org/10.1177/25152459211007467

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why Hypothesis Testers Should Spend Less Time Testing Hypotheses. *Perspectives on Psychological Science*, *16*(4), 744–755. https://doi.org/10.1177/1745691620966795

Szollosi, A., & Donkin, C. (2021). Arrested theory development: The misguided distinction between exploratory and confirmatory research. *Perspectives on Psychological Science*, *16*(4), 717–724. https://doi.org/10.1177/1745691620966796

Table 1

Criterion	Original	Theory	Action
F1	(Meta)data are assigned a globally unique and persistent identifier	Theory and its associated metadata has a global, unique and persistent identifier for each version (using semantic versioning)	Rephrased
F2	Data are described with rich metadata	Theory is described with rich metadata	~same
F3	Metadata clearly and explicitly include the identifier of the data it describes	Metadata clearly and explicitly include identifiers for all the versions of the theory it describes	~same

Table 1 continued

Criterion	Original	Theory	Action
F4	(Meta)data are registered or indexed in a searchable resource	Theory and its associated metadata are included in a searchable repository	NEEDS WORK! GitHub is indexed by Google I believe, but ideally, we'd like our theories to show up in Google Scholar, or even dedicated academic search enginges. Where could we put them to realize this?
A1	(Meta)data are retrievable by their identifier using a standardized communications protocol	Theory and its associated metadata are accessible by their identifier using a standardized communications protocol	~same

Table 1 continued

Criterion	Original	Theory	Action
A1.1	The protocol is open, free, and universally implementable	The protocol is open, free, and universally implementable	Same
A1.2	The protocol allows for an authentication and authorization procedure, where necessary	The protocol allows for an authentication and authorization procedure, where necessary	Do we need this?
A2	Metadata are accessible, even when the data are no longer available	Theory metadata are accessible, even when the theory is no longer available	Do we need this?
I1	(Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation	Theory and its associated metadata use a formal, accessible, shared and broadly applicable language to facilitate machine readability and reuse	Rephrased

Table 1 continued

Criterion	Original	Theory	Action
I2	(Meta)data use vocabularies that follow FAIR principles	(Meta)data use vocabularies that follow FAIR principles	NEEDS WORK! I think this is where we explain the value of e.g. universal graph languages like Aaron and Max'
I2S.1	-		
I2S.2	-		
I3	(Meta)data include qualified references to other (meta)data	(Meta)data includes qualified references to other (meta)data, including previous versions of the theory	Rephrased. I envision a LinkList-like structure where each theory version references its ancestor
I4S	-		Discard

Table 1 continued

Criterion	Original	Theory	Action
R1	(Meta)data are richly described with a plurality of accurate and relevant attributes	Theory and its associated metadata are richly described with a plurality of accurate and relevant attributes	Needs work. How do we envision this? Keywords? ISBN-like codes for variable type?
R1.1	(Meta)data are released with a clear and accessible data usage license	(Meta)data are released with a clear and accessible license	Needs work: which license is good for theory?

Table 1 continued

Criterion	Original	Theory	Action
R1.2	(Meta)data are associated with detailed provenance	(Meta)data are associated with detailed provenance	Relates to I3; I think this point relates more to the theory's ancestry, and I3 relates to e.g. incorporating other theories within a theory (e.g., theory of measurement inside of structural theory)

Table 1 continued

Criterion	Original	Theory	Action
R1.3	(Meta)data meet domain-relevant community standards	Theory metadata and documentation meet domain-relevant community standards	These standards do not yet exist; we can take a first step towards developing them and recommend that this be an active area of development

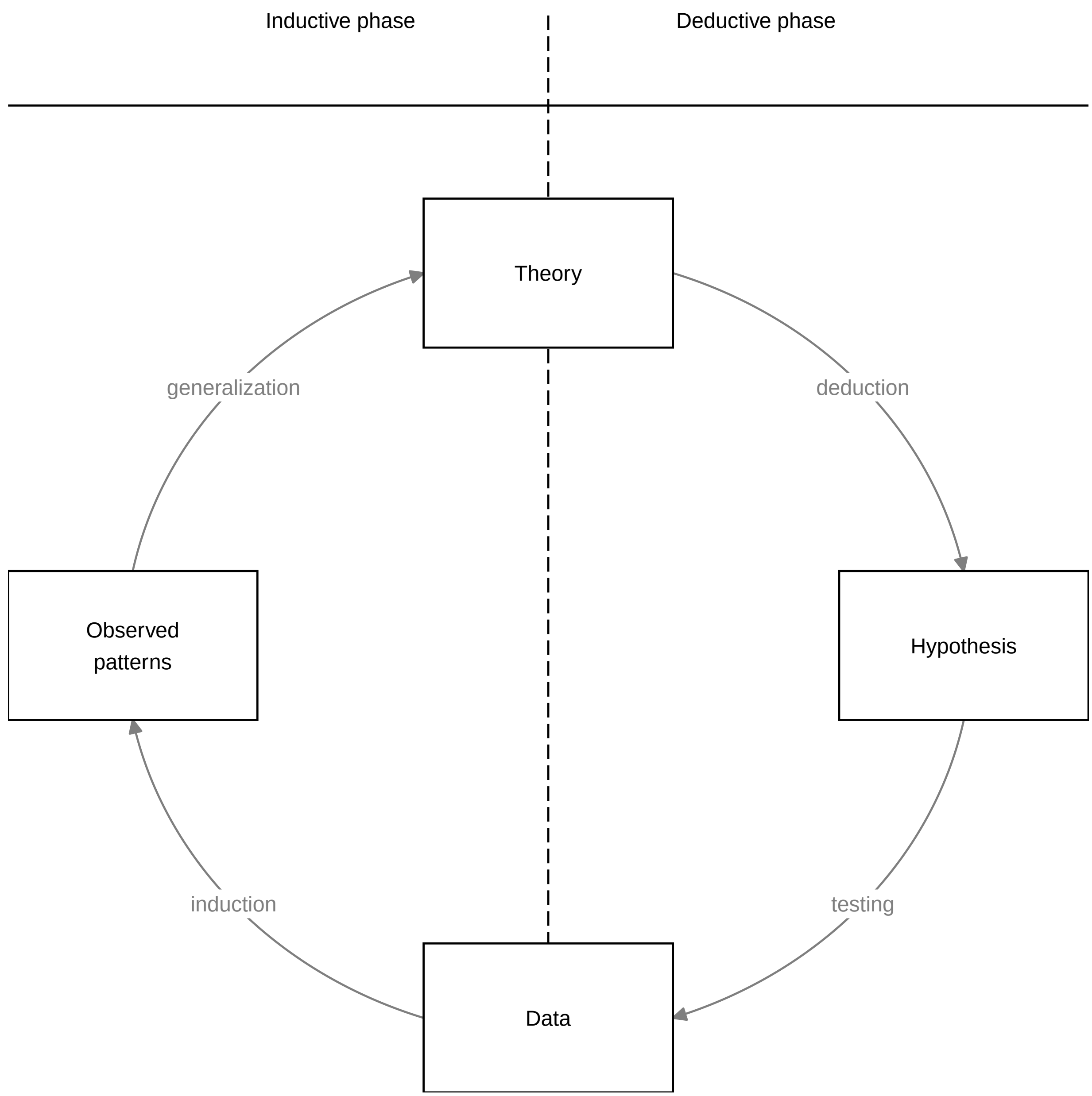

Figure 1. A take on the empirical cycle by De Groot