Methodological empiricism and the choice of measurement models in social sciences*

Clayton Peterson

Abstract

Realism is generally assumed as the correct position with regards to psychological research and the measurement of psychological attributes in psychometrics. Borsboom et al. (2003), for instance, argued that the choice of a reflective measurement model necessarily implies a commitment to the existence of psychological constructs as well as a commitment to the belief that empirical testing of measurement models can justify their correspondence with real causal structures. Hood (2013) deemphasized Borsboom et al.'s position and argued that the choice of a reflective measurement model does not necessarily require ontological commitments, though, in his view, it does necessitate a commitment to minimal epistemic realism. Although these arguments are formulated with regard to psychological research, they can actually be generalized to other disciplines in social sciences that use similar methodologies and statistical techniques. In Hood's opinion, empiricism does not suffice to provide an adequate account of the choice of reflective measurement models given that this choice requires an appeal to causal explanations. In this paper, we argue against Hood and answer this challenge, providing epistemic foundations for social science research that do not appeal to realism.

---

*Forthcoming in the European Journal for Philosophy of Science.

1

Acknowledgments

1   Introduction

Realism is an important part of the literature in the philosophy of human and social sciences, specifically in psychology. Besides various positions one finds with respect to folk psychology, representations, and mental states (e.g. Fodor 1980; Jackson and Pettit 1990; Wright 2002), realist arguments can also be found in the literature on measurement and psychometrics (cf. Michell 1997, 2003, 2004, 2005, 2013; Borsboom et al. 2003; Borsboom 2005, 2008; Hood 2013).

The predominance of realism with respect to psychological measurement can be understood in light of the historical context underlying the methodological guidelines suggested for the development of measurement instruments. Indeed, realism has been developed in reaction to the operationalist understanding of measurement models (cf. Waters and Pennington 1938; Stevens 1939; Bergmann and Spence 1941; Israel and Goldstein 1944; Cronbach and Meehl 1955). The operationalist thesis amounts to the claim that latent traits, represented by latent constructs, only exist through the equations defining these constructs. One problem with operationalism, however, is that the idea that latent traits do not exist beyond their statistical definition is conflicting with the rigorous methodology used to define said traits, which seems to imply that there are latent traits being measured. Put differently, it appears that operationalism is not sufficient to provide a rational justification of the methodological choices that are made during the development of measurement instruments. As a result, realism has been presented as an alternative epistemic standpoint that can rationally explain methodological decisions in psychological research and, more generally, in social sciences.

In the last decades, there has been a tendency to assume that a realist

understanding of psychological attributes and measurement models is the only viable option, and the burden of the proof has been left to anti-realists to provide an intelligible conception of the methodology used in social sciences. With respect to psychological research, the culmination of the realist position can be found in the work of Borsboom et al. (2003), who argue that psychometrics necessarily requires epistemological and ontological realist standpoints. However, Hood (2013) deemphasized that position and argued that psychometrics only necessitates epistemological realism and, in this respect, it does not necessarily require ontological commitments.

Hood's main thesis is dubbed *methodological realism*, which in turn relies on the principle of *minimal epistemic realism*.

Methodological realism The rationality of certain methodological decisions requires a commitment to minimal epistemic realism (Hood 2013, p.741).

Minimal epistemic realism It is, in principle, possible to justify hypotheses that posit unobservable entities (Hood 2013, p.740).

This position is grounded on Leplin's (1986; 1997) analysis, who argues against instrumentalism that realist assumptions are required to justify and explain methodological decisions in science. Hood (2013, pp.742,759), building on Leplin's thesis, argues that this position also holds in psychology: Methodological choices in psychological research and psychometrics cannot be conceived as rational under an anti-realist or an empiricist epistemic position (see also Leplin 1986, p.32).

Hood's theses can be summarized as follows. First, a rational (or non-arbitrary) choice to model covariation reflectively rather than formatively necessitates that researchers are committed to the idea that, in principle, it is possible to justify the hypothesis that posit the existence of latent traits. Second, trying to estimate (latent) abilities necessitates that researcher are committed to the idea that, in principle, it is possible to justify the hypothesis that there are such (latent) abilities and, further, they are committed to the existence of such abilities. Third, a rational (or non-arbitrary) choice to use a longitudinal research design rather than a cross-sectional one to model within-subject variation necessitates that researchers are committed to the idea that, in principle, it is possible to justify the hypotheses stipulating the existence of latent traits.

Although Hood concentrate on the (alleged) necessity of realism to justify methodological decisions in psychological research, his arguments can be generalized to other disciplines in social sciences that use similar methodologies and statistical techniques, such as sociology, education, or management. Thus generalized, Hood's position amounts to the claim that anti-realists and empiricists cannot provide a rational account of methodological choices made in social sciences with respect to latent attributes and measurement models.

In what follows, we answer Hood's challenge and articulate an empiricist position that can rationally explain methodological decisions made in social sciences. Given that Hood's arguments are aiming at psychological research specifically, we will concentrate on that as well, keeping in mind that our analysis can be generalized to other disciplines that use similar statistical techniques. In response to Hood's thesis, we argue that minimal epistemic realism is not a necessary epistemic standpoint required to justify methodological decisions made with respect to measurement models insofar as it cannot be defended consistently when there is underdetermination. The objective is not to establish or deny that researchers are committed to minimal epistemic realism but is rather to deny that minimal epistemic realism is *necessary* to make sense of the non-arbitrariness of psychological research. Our arguments are grounded on specific examples taken from the scientific literature, supporting the idea that only *methodological empiricism* is necessary to justify methodological decisions in psychology and, more generally, in social sciences.

We begin in section 2 with a summary of Hood's arguments in favor of methodological realism as well as Borsboom et al.'s position, on which Hood's position partially relies. The thesis defended within this paper, *minimal epistemic empiricism*, is presented in section 3, and Hood's arguments are discussed in section 4. In a nutshell, we argue that underdetermination implies that minimal epistemic realism leads to inconsistent beliefs with regard to competing statistical models. We conclude in section 5 with a summary of our analysis.

2 Minimal epistemic realism

2.1 Measurement models

The core of Hood's argument regards the (methodological) choice of modeling latent attribute reflectively rather than formatively. In social sciences, latent traits (understood broadly as latent characteristics), such as personality traits (e.g., John and Srivastava 1999) or the satisfaction of psychological needs (e.g., Deci and Ryan 2008), are formally represented by latent constructs, otherwise known as latent variables or factors. While empirical phenomena can be observed, latent traits cannot. Accordingly, claims regarding latent traits can only be established through statistical inference.

Latent constructs, which are abstract representations of latent traits, are defined through measurement models. They are measured indirectly via participants' responses to items, which are considered as empirical indicators of these constructs. An item is a statement within a questionnaire, to which participants attribute scores. There are various types of items and scales (see for example Mellenbergh 1994 or Stevens 1946), and a well-known example would be an item scored on a Likert-type scale ranging from 1 = *completely disagree* to 5 = *completely agree.*

As there is a distinction to be made between a *trait* (characteristic) and a *construct* (latent variable, factor), there is a distinction between *subject reports* (e.g., report of one's age) and *observable traits* (e.g., one's age). Although researchers are trying to establish claims regarding observable and latent traits, they are rather dealing with subject reports and latent constructs.

A measurement model can take either one of two forms (cf. Edwards and Bagozzi 2000). A *formative* measurement model considers that the variation on the latent variable is a consequence of the between-subjects variation in the items.[1] Mathematically speaking, the latent variable is modeled as the dependent variable, whereas subject reports on items are the independent variables. Hence, the model is such that subject reports predict the latent variable. Such a model is then considered as evidence supporting the idea that a variation on observable traits causes a variation on the latent trait. A good example of a latent trait modeled formatively would be the socio-economic status (cf. White 1982). It is defined on the grounds of variables

---

[1]See Borsboom et al. (2003) for a discussion of the distinction between inter-subjects and intra-subject variation.

including the level of education and income. Formally, a variation on the construct is modeled as a consequence of the variation on subject reports on level of education or income. This mathematical abstraction is meant to represent the idea that different levels of education or income imply different socio-economic status. That is, the model is meant to represent the idea that socio-economic status depends on level of education and income. As such, the trait *socio-economic status* is conceptualized as an effect of the traits *level of education* and *income*, and this is depicted within a mathematical model where subject reports of level of education and income predict the latent variable interpreted as the socio-economic status.

Formative models, however, are not the only kind of measurement models that are used in psychology. Indeed, some measurement models are meant to *explain* the between-subjects variation on items. As an example, consider a researcher that thinks that some psychological phenomena could be explained using the notion of the satisfaction of psychological needs, say, the satisfaction of the need for autonomy at work (e.g., Van den Broeck et al. 2010). Assuming that there is such a thing, how could one measure the satisfaction of the need for autonomy? The researcher might decide to start with a Likert-type scale in which 1 represents strongly disagree, 2 disagree, 3 neither agree nor disagree, 4 agree, and 5 strongly agree. Then, the researcher will assume that if there is such a latent trait, then one would expect to be able to observe empirical consequences of this trait. For instance, the researcher might assume that if a participant's need for autonomy is satisfied at work, then one will score high (i.e., 4 or 5) on an item such as "I feel like I can be myself at my job".

This is the rationale underlying a *reflective* measurement model. Researchers postulate that there are latent psychological attributes and try to indirectly observe them by postulating their plausible observable consequences. In this case, the observable consequence is the between-subjects variation on the items, and the latent attribute is modeled as a latent variable. A measurement model is said to be reflective when the latent variables are taken as explanations of the between-subjects variation on the items. When modeled reflectively, the latent trait is (informally) understood as a cause that explains the between-subjects variation on the items (cf. Spirtes et al. 1991, 2000 and Glymour 2001). Items are thus considered as indicators of latent traits insofar as variation on these items is an observable effect of these hypothesized causes. Mathematically, and in contrast to formative measurement models, latent traits are modeled as independent variables,

6

whereas items are conceived as dependent variables. Consequently, it is the latent variable that predicts the variation on the items.

Latent constructs and measurement models are identified through the use of exploratory (EFA) and/or confirmatory (CFA) factor analysis (see Ullman 2013 or Kline 2013; for a more philosophically oriented presentation, see also Johnson 2016).[2] As such, measurement models are special cases of structural equation models. Without going into the formal details (see Peterson 2018 for further explanations), the basic idea is to model an item as a linear combination of some factors (i.e., latent variables) in addition to an error term, which accounts for possible error in measurement. Items are thus considered as dependent variables which (between-subjects) variation is due to (i) some independent variables, namely the factors and (ii) error. Hence, factors are considered as mathematical predictors that explain the between-subjects variation on the items.

2.2 Borsboom et al.'s position

The starting point of Hood's argument is Borsboom's (2005; defended previously by Borsboom et al. 2003) position that the choice of a reflective model necessitates an ontological commitment to the existence of psychological attributes and, further, a commitment to the belief that empirical testing of measurement models can justify their correspondence with real causal structures (i.e., can justify that measurement models are true). Hood's summary of Borsboom's position is as follows:

> The choice to model covariation with a reflective model, as opposed to a formative model, requires a commitment to realism. There is no a priori reason to prefer one model to the other unless one harbors realist commitments. Without such commitments, the decision seems arbitrary. Indeed, it would seem irrational to prefer the model that carries with it ontological commitment if one does not embrace some form of realism (Hood 2013, p.749).

---

[2]Michell (1997, 2005) argues that measurable attributes must have a quantitative structure and that, since the structure of the measurement unit must be isomorphic to the structure of the attribute, it follows that psychological attributes, if they are quantitative at all, cannot be measured by Likert-type scales. See Borsboom and Mellenbergh (2004) and Hood (2013, 2008) for an analysis of Michell's position.

Borsboom et al. (2003) provide three arguments to motivate that position. The first argument regards the evaluation of a participant's position on a latent variable. In their view, realism is entailed by the fact that there is a probability of error when estimating a participant's position on a latent variable: If there is a possibility of mistake, then realism holds since "there must be something to be wrong about (Borsboom et al. 2003, p.209)". In this respect, Borsboom et al.'s position is that measurement necessarily implies an ontological commitment insofar as one can only measure things that exist.

The second argument relies on the notion of estimation of parameters and follows the same line of reasoning. In their view, estimation is a realist concept insofar as only things that are real can be estimated. The argument, however, goes a step further: Estimation of parameters means estimating the *true* value of a parameter.[3] When estimating a parameter, one is most likely to be wrong about it. However, "it is impossible to be wrong if it is not possible to be right (Borsboom et al. 2003, p.209)". That being said, being right or wrong is accomplished with respect to the *true* value of a parameter.[4] Hence, to their mind, estimation requires (i) that there is something to be estimated (i.e., a realist ontological commitment) and (ii) that estimations of parameters are true or false (i.e., epistemological and semantical realist commitments).

Finally, the third argument regards the assumption that there is a *true* model, which, according to the aforementioned reasoning, requires a (full-blown) realist perspective. When researchers develop a measurement model to represent latent traits, they assume that it is true and then test it. That is, they assume that the measurement model is adequate, meaning that it corresponds to an actual state of affairs (Borsboom et al. 2003, p.210). Consequently, in their view, assuming a measurement model that is tested necessitates realist commitments.

---

[3]This is an assumption underlying frequentist factor analysis techniques (e.g., using maximum likelihood extraction). It should be noted, however, that there are Bayesian foundations for factor analysis that rather use a probability distribution (e.g., Muthén and Asparouhov 2012).

[4]In their view, realism implies that truth should be understood as correspondence (Borsboom et al. 2003, p.209). It should be noted, however, that metaphysical realism does not necessarily imply correspondence truth and, similarly, that correspondence truth does not necessarily imply metaphysical realism. It can be argued that the semantical and the ontological issues are independent (e.g., Horwich 1996).

2.3 Hood's reply

Hood's (2013) criticism of Borsboom et al.'s (2003) position is compelling. As it stands, Borsboom et al.'s (2003) analysis could benefit of more conceptual distinctions. Indeed, they treat realism as if it were one unambiguous philosophical position that necessarily entails specific ontological, semantical and epistemological claims. In this respect, they fail to address that there are different aspects of realism. Hood (2013) addresses this issue and starts his analysis by distinguishing two different aspects of Borsboom et al.'s (2003) position, namely their *ontological* and their *epistemological* realist commitments: On the one hand, they are committed to the existence of latent traits and, on the other, they are committed to the idea that researchers are justified to believe in successful theories.

The core of Borsboom et al.'s (2003) argument is that the choice of a reflective measurement model necessitates a commitment to realism. Hood (2013, pp.748-51), however, argues that the choice of a reflective measurement model only requires minimal epistemic realism but does not require any ontological commitment. His (correct) argument refuting that psychometrics requires ontological commitments can be summarized as follows: Measurement models are hypotheses that can be tested, in which case they

> may enjoy more or less empirical support. [...] Proposing a hypothesis does not *ipso facto* commit one to the existence of the entities postulated therein. That psychometricians model covariation reflectively while remaining agnostic about the ontological status of their latent variables is perfectly intelligible (Hood 2013, p.750).

Hood is perfectly right on this point. This argument addresses Borsboom et al.'s position that the choice of a reflective measurement model necessitates ontological commitments and, further, it also addresses their third argument that postulating a *true* model necessitates similar engagements.

However, Hood goes a step further, which, in our opinion, is a step too far. Even though he rejects the necessity of the ontological claim, he argues in favor of methodological realism by defending that psychometrics requires a commitment to minimal epistemic realism. His argument is threefold.

2.3.1 Choosing to model reflectively

First, Hood argues that the choice between a formative and a reflective measurement model requires a commitment to minimal epistemic realism. For instance, he considers a practical study where the measurement model is empirically tested through various EFA and CFA techniques (cf. Jensen and Weng 1994). The aim of Jensen's and Weng's (1994) study is to establish the robustness of the intelligence factor g using various methods. Hood (2013, p.750) begins by pointing out that "reflective models are hypotheses about the structure of covariation", which are also "causal hypotheses about the underlying, i.e., latent, common cause(s) of test behavior" and, as such, "they are testable against novel data through confirmatory factor analysis". Hood's argument runs as follows:

> It would seem to make little sense to be "testing" these measurement models against data if Jensen did not think that there was the possibility of confirming the g hypothesis. [...] Merely subjecting such models to a confirmatory factor analysis suggests that one is aiming to confirm the model, causal relationships, unobservables, and all. Thus, the selection of models that posit latent sources of variation and the subsequent testing of those models, as opposed to models that do not make such empirical claims about the latent structure of ability give good grounds for methodological realism (Hood 2013, p.750).

Quoting Jensen and Weng (1994, p.254), Hood (2013, p.751) argues that minimal epistemic realism is necessary to make sense of the idea that

> factor analysis would have little value in research on the nature of abilities if the discovered factors did not correspond to real elements or processes, that is, unless they had some reality beyond the mathematical manipulations that derived them from a correlation matrix.

In other words, measurement models are hypotheses regarding the relationship between latent and observable traits. Hence, by choosing measurement models that postulate causal relationships and by submitting them to empirical verification via CFA, researchers need to minimally assume that it is, in principle, possible to justify these hypotheses (which specify the existence of these latent traits) via empirical data.

2.3.2 Prediction of abilities

Second, Hood (2013, pp.754-5) argues that the prediction of abilities in psychometrics does not only require a commitment to minimal epistemic realism, but also requires an ontological commitment to the existence of the attribute. This argument is in line with Borsboom's et al. (2003) second argument, specifying that estimation of parameters requires that (i) there exists a parameter that is estimated and (ii) estimations are true or false with respect to the true value of the parameter. In Hood's (2013, p.755) terms:

> If psychometricians did not believe that the attribute in question existed, and if they did not further believe that claims about individual position on the attribute were corrigible, estimation procedures and attempts to reduce sampling error by giving a diverse battery of tests would be a perverse and arbitrary enterprise. Consequently, estimating ability requires both methodological realism and a commitment to the existence of latent traits.

2.3.3 Longitudinal studies and within-subject variation

The third argument regards the choice to perform longitudinal studies rather than cross-sectional ones. Specifically, it concerns the choice to model *within*-subject variation rather than *between*-subjects variation on items.[5] While a cross-sectional study uses data gathered at only one time point, longitudinal studies use data gathered at subsequent time points. The aforementioned measurement models, obtained using EFA and/or CFA, take place within the context of cross-sectional studies. These models provide an explanation of the *between*-subjects variation on the items.

Some researchers, however, are interested in what is happening at the level of individuals. That is, they are interested in the *within*-subject variation on the items, not the between-subjects variation. To see what happens at the level of the individual, one needs many measurements at different time points. Hood (2013, pp.755-8) argues that the (methodological) choice to model within-subject variation rather than between-subject variation, which he dubs the *idiographic turn*, requires a commitment to minimal epistemic

---

[5]Longitudinal studies is a rather broad category and it does not necessarily model within-subject variation. There are longitudinal research designs that concentrate on between-subjects variation.

realism. For Hood, choosing a longitudinal design that can be used to explain what happens at the level of the individual indicates that researchers assume that the data gathered will provide better evidence to justify the existence of the traits, which amounts to assert minimal epistemic realism.

Nonetheless, the idiographic turn does not require any ontological commitment, given that postulating hypotheses that posit the existence of latent traits does not engage researchers to be committed to the existence of these traits.

3   Minimal epistemic commitments

3.1   Methodological empiricism

In its current formulation, minimal epistemic realism is somewhat ambiguous. It stipulates that it is (in principle) possible to justify *hypotheses* that postulate the existence of unobservable entities. Implicit to Hood's formulation is the idea that empirical data can justify hypotheses (in our case, statistical models). But what does it mean to say that data can justify a hypothesis? Although this formulation is widely used in the common language, it is actually misleading, notwithstanding that it will have a different meaning for a realist and an anti-realist. Given that hypotheses are declarative sentences (i.e., which have the potential to be true of false), realists will tend to understand 'justifying a hypothesis' as 'empirical data can justify that hypotheses are true'. But this would be misleading given that empirical data do not justify that hypotheses are true but rather *confirm* hypotheses. As we know, confirmation of a hypothesis through empirical data does not provide a justification that said hypothesis is true. For an anti-realist, to say that 'data justify a hypothesis' will rather mean that data provide support in favor of the hypothesis. As such, the empirical data give us grounds to believe that the hypothesis is more or less probable.

Empirical data does provide a justification, but it is not hypotheses that are subject to justification per se. Rather, *epistemic states*, such as *knowledge* and *belief*, are subject to justification by empirical data. Although one might say that data provide support in favor of a hypothesis, this only means that this empirical data can be presented as a reason that justifies a belief in said hypothesis. To avoid any ambiguity, we propose to rephrase the thesis of minimal epistemic realism as follows.

Minimal epistemic realism It is, in principle, possible to justify the belief in hypotheses that posit unobservable entities.

Before going further, it is worth discussing how *justification* is understood within the context of minimal epistemic realism. The literature that followed Gettier's (1963) counterexamples to knowledge understood as justified true belief provided us with different ways to understand how beliefs are justified. The weakest form of justification happens when one has reasons or evidence to support a belief. However, Gettier's counterexamples showed that this type of justification is insufficient: Merely having reasons to believe does not imply that one is indeed justified to believe, for one may believe something that is actually true but for the wrong reasons. To put it differently, one might be justified (in this sense) to believe, although the belief, in itself, is not justified (cf. Ball 2012; Feldman 2004). This example brings to light that there is a distinction to be made between 'a person is justified to believe' and 'a belief is justified'. While the former implies a doxastic aspect, the latter is rather a form of *propositional justification* (see also Turri 2010). To avoid any ambiguity, let us refer to the first type of justification as *weak doxastic justification*. It happens when one possesses evidence that support, to some degree, a belief, although the belief might turn out to be false after all. In contrast, *propositional justification* happens when a belief is supported by sufficient reasons. In other words, a belief is said to be propositionally justified when the reasons that support it are sufficient to guarantee its truth. As such, when a belief is propositionally justified, it is supported by the 'right' evidence (cf. Ball 2012; Turri 2010). It is noteworthy, however, that propositional justification does not imply that one will believe on the grounds of these sufficient reasons. Indeed, a belief can be propositionally justified, although one believes it but not for these reasons, or does not believe it at all. Nonetheless, if a belief is propositionally justified and one believes it on these grounds, then one will be *strongly doxastically justified* to believe. That is, one will believe for the right reasons.

In light of these distinctions, one might wonder which type of justification Hood had in mind when arguing in favor of minimal epistemic realism. As it happens, the notion of justification that is implicit to minimal epistemic realism lies in between weak doxastic justification and propositional justification. Indeed, Hood argues that scientists need to assume that the way they justify their beliefs might serve as propositional justification, although it might as well be that scientists are only weakly doxastically justified to believe in

latent constructs (see Hood 2013, p. 742). In other words, scientists need to think that it is at least *possible* that their methodology provides them with the right reasons to believe, otherwise it would be pointless to use that methodology. Nonetheless, it is still possible that latent constructs do not exist, so there is a possibility that the methodology is insufficient to justify scientists' beliefs. Given that Hood insists on the idea that it is possible that the methodology used in human and social sciences provides scientists with propositional justification, this is how we will understand *justification* in the remaining of this paper.

Belief is understood minimally as a KD-modality (cf. Halpern et al. 2009). Thus conceived, a person cannot believe in both a proposition and its negation (i.e., it is false that one believes in both φ and $\neg\varphi$). That is, beliefs must be consistent. In contrast to minimal epistemic realism and following van Fraassen (1980), we propose to adopt the following thesis.

Minimal epistemic empiricism   It is, in principle, possible to justify the belief in the empirical adequacy of hypotheses that posit unobservable entities.

However, we should keep in mind that we are trying to provide an epistemic standpoint that can justify methodological choices rather than beliefs per se. Methodological choices involve pragmatic elements. In this context, one should distinguish between *acceptance* of a theory (or model) and *belief*. Acceptance of a theory implies using said theory to explain and predict empirical phenomena. The notion of acceptance captures this pragmatic aspect, whereas *belief* does not necessarily. In the aforementioned formulation of minimal epistemic empiricism, 'justifying the belief in a hypothesis' means *accepting* the hypothesis. As such, minimal epistemic empiricism might be rephrased as follows.

Minimal epistemic empiricism   Accepting hypotheses that posit unobservable entities requires a belief in their empirical adequacy.

Methodological empiricism   The rationality of certain methodological decisions requires a commitment to minimal epistemic empiricism.

By 'empirically adequate', we mean that hypotheses (theories, models, explanations, etc.) need to be consistent with our observations. That is, they

need to be "correct, or approximately correct enough (Bhakthavatsalam and Cartwright 2017, p.446)". In the case of EFA and CFA (and, more broadly, structural equation models), the consistency of the model with the data (i.e., its empirical adequacy) is evaluated through various goodness-of-fit criteria, which will be discussed further in section 4.1 below.

3.2   Necessary epistemic attitude

From now on, we will argue that it is the thesis of methodological empiricism that is adequate for social science research and, further, that it is the principle of minimal epistemic empiricism that is the *minimal* assumption required to make sense of the non-arbitrariness of methodological decisions in social sciences.

Before going further, it is important to emphasize that endorsing minimal epistemic realism implies endorsing minimal epistemic empiricism. Empirical adequacy is, among others, one goal of the scientific enterprise. If a hypothesis h is not empirically adequate, then a belief in h is not warranted. Put differently, if a belief in h is justified, then h is empirically adequate. Similarly, if a hypothesis is true, then it is empirically adequate. Now, consider the following principle, which is comparable to Hempel's (1943) special consequence condition, stipulating that if some empirical evidence confirms h, then it also confirms every consequences of h.

Transitivity of justification   If a belief in φ is justified, then a belief in the consequences of φ is also justified.

Assuming that beliefs are closed under consequence, we can reason as follows: From the premise 'if a hypothesis is true, then it is empirically adequate', we obtain 'the belief that a hypothesis is true implies the belief that it is empirically adequate'. From the transitivity of justification, we thus obtain that 'if the belief that a hypothesis is true is justified, then the belief that it is empirically adequate is justified'. Hence, if one assumes that it is possible to justify a belief in a hypothesis, then one also assumes that it is possible to justify a belief in its empirical adequacy.

The minimal assumption that is required by disciplines that use factor analyses to be non-arbitrary is minimal epistemic empiricism. Put differently, researchers in social sciences only need to be committed to methodological empiricism. Researchers need to assume that it is in principle possible to

justify the belief in the empirical adequacy of reflective measurement models if the empirical tests of these models are taken as a justification of said belief. It would be pointless to empirically test measurement models if such tests could not provide any form of justification whatsoever.

This is quite straightforward and is established by *reductio*: Assume (i) that a researcher considers that empirical testing of a measurement model is taken as a justification of the belief that the measurement model is empirically adequate but (ii) that it is not possible to justify the belief in the empirical adequacy of a reflective measurement model. Assuming that possibility and necessity behave as *KT*-modalities (cf. Chellas 1980), this is inconsistent.[6]

That minimal epistemic empiricism is a necessary condition for social science research can also be seen in light of the rationale behind scientific reasoning and hypothesis testing. Researchers postulate that some traits can be used to explain some phenomenon. This does not imply a commitment to the existence of these traits. Researchers then try to determine how said traits could be identified. Assuming that there might be such things as latent traits, one will postulate plausible empirical consequences that such traits would have. For instance, one might postulate that if a participant's need for autonomy is satisfied at work, then one will score high on an item such as "I feel like I can be myself at my job". The rationale behind that conception is that if a participant does indeed score high on this item, then it will confirm the hypothesis. That is, it will provide support in favor of the belief in the empirical adequacy of the explanation expressed by the reflective measurement model. Hence, minimal epistemic empiricism is implicit to the rationale behind hypothesis testing and confirmation. To paraphrase van Fraassen (1980, p.12), social science research aims to give us measurement models which are empirically adequate; and acceptance of a measurement model involves as belief only that it is empirically adequate.

Hood argues that postulating the existence of latent traits does not commit researchers to the existence of these traits. Put differently, assuming that latent traits might exist does not imply that one is committed to the idea that they actually do exist. The same argument can be used to argue against minimal epistemic realism. One can be committed to the idea that it is possible to justify a belief in a hypothesis's empirical adequacy while remaining agnostic regarding whether or not it is possible to justify a belief in a hypoth-

---

[6](i) means that the researcher considers that 'p' while (ii) means that she considers '$\neg \Diamond p$'.

esis's truth. That is, one can be committed to minimal epistemic empiricism while remaining agnostic with respect to minimal epistemic realism.

As a caveat, it should be noted that we are not proposing that empirical adequacy is a good substitute for truth.[7] Similarly, we are not suggesting that minimal epistemic empiricism is the actual attitude of researchers. Following Hood's argumentation, we are only proposing that believing in the possibility of justifying a belief in the empirical adequacy of a hypothesis is a necessary condition for the non-arbitrariness of psychological research and social science research in general. That is, the only requirement to motivate acceptance of reflective measurement models is the belief in their empirical adequacy.

4 Answering Hood's arguments

4.1 Model selection

Choosing to model reflectively because it allows explaining the between-subjects variation on the items does not require a commitment to the belief that the data can be used to justify the belief that a hypothesis is true. The choice of a reflective measurement model is rational insofar as the latent trait, modeled as a factor, is taken as an explanation of the variation on the items, and the data gathered provide grounds to justify the belief in the empirical adequacy of this hypothesis.

One cannot consistently defend that minimal epistemic realism is necessary to justify the choice of a reflective measurement model. Assume that a researcher is necessarily committed to the belief that data can justify the belief in a hypothesis's truth. An important aspect of social science research is model selection. Various (reflective) measurement models can be used to model the data. To exemplify our point, let us consider an actual measurement model taken from the literature (cf. Trépanier et al. 2013). This model is well-confirmed and is theoretically grounded.

When conducting an EFA, researchers have to choose between models with different numbers of factors, notwithstanding that they also have to choose which items will be kept in the final measurement model. Consider a measurement model aimed at depicting psychological need satisfaction at work (cf. table 1), a concept proposed by Self-determination theory (Deci

---

[7]This have been contested. See, for instance, Leplin (1987, 1997) or Psillos (1999).

	Items
aut1	I often feel like I have to do what I am ordered to.
aut2	If I could, I would do things differently.
aut3	My assignments correspond to what I really want to do.
aut4	I feel free to accomplish my work as I think it should be done.
com1	I feel competent at my job.
com2	I am good at my job.
com3	I understand the tasks that should be accomplished.
com4	I feel like I can accomplish all tasks at work, even the tough ones.
rel1	I don't feel related to my coworkers.
rel2	I feel like I am part of a group at work.
rel3	I do not interact with others at work.
rel4	I can talk about subjects I find really important with others at work.

Table 1: Items for need satisfaction at work (Van den Broeck et al. 2010)

and Ryan 2008). According to this theory, individuals have three innate psychological needs that must be satisfied to achieve optimal functioning: need for autonomy (experience of volition and self-endorsement of one's actions), need for competence (perception of being able to master one's environment and attaining important outcomes within it), and need for relatedness (sense that one has significant and meaningful interpersonal relationships).[8] These three characteristics are captured through various items, as presented in table 1.[9] These items are scored on a five-points Likert-type scale ranging from 1 (totally disagree) to 5 (totally agree).

When performing an EFA using these items on a data set with a sample size of 1200, we obtain the factor loadings presented in table 2.[10] The usual recommendation is to consider factor loadings that are greater than .3 (cf. Field 2009). That is, only items with factor loading greater than .3 should be considered as indicators of the factor. However, some authors have suggested that other levels of factor loadings can also be considered significant, depending on the sample size. For instance, with a sample size of 1200 (as in our example), one might consider items with factor loadings that are strictly greater than .162 (cf. Field 2009, p.644).

---

[8]See Trépanier et al. (2013, p.126).

[9]The scale used in Trépanier et al. (2013) actually contains more items. For the sake of this example, we were only provided with part of the scale.

[10]The analysis was performed using a maximum likelihood extraction as well as a varimax rotation. Loadings greater than .3 are in bold, and loadings lesser than .162 have been omitted.

	Competence	Relatedness	Autonomy
com1	**.844**		
com2	**.891**		
com3	**.557**	.172	.183
com4	**.595**		.222
aut1			**.496**
aut2		.162	**.679**
aut3	.191	.167	**.648**
aut4	.190		**.669**
rel1		**.671**	
rel2		**.790**	.206
rel3		**.650**	
rel4	.176	**.554**	.252

Table 2: EFA for need satisfaction

As we can see, one will obtain two different measurement models depending on the guideline that is followed. If we follow the usual recommendation and consider only items with a loading greater than .3, then the factorial structure is quite clear: com1-com4 are indicators of satisfaction of the need for competence; aut1-aut4 are indicators of satisfaction of the need for autonomy; and rel1-rel4 are indicators of satisfaction of the need for relatedness. This is a three-factor model containing all items (hypothesis h). However, if we follow the alternative recommendation and consider all loadings strictly greater than .162, then we will need to exclude com3-4, aut3-4, and rel2-4 given that items need to load significantly on only one factor (i.e., one wishes to obtain a simple structure, cf. Thurstone 1947, 1954). This is also a three-factor model but with only two items per factor (hypothesis h').[11] In light of these divergent recommendations, there are two different measurement models to choose from and, as such, there is underdetermination.

Now, consider the case of confirmatory factor analysis. In CFA, the researcher has to specify *a priori* the measurement model that is to be tested. She has to specify the number of factors but she also has to specify which item should load on which factor. These choices have to be justified in light of competing plausible measurement models. Given the aforementioned example, one might want to test both h and h' in order to see which one is the best fitting model. The empirical adequacy of a (measurement) model in CFA

---

[11]It can be argued that at least three items per factor are required in an adequate measurement model (cf. Velicer and Jackson 1990). We will consider only two items per factor for the sake of our example.

	CFI	TLI	RMSEA	SRMR
h	.946	.930	.058	.049
h'	.998	.996	.019	.010

Table 3: Goodness-of-fit indexes obtained by CFA

is based on various goodness-of-fit criteria (cf. Kline 2005 or Byrne 2012). For example, there is the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR).[12] CFI, TLI, SRMR and RMSEA take a value between 0 and 1, and the empirical adequacy of the model can be evaluated using these fit indices. While CFI and TLI need to be as close to 1 as possible ($x \geq .9$), RMSEA and SRMR need to be closest to 0 ($x \leq .08$).

The results of the CFA performed on the aforementioned data set are presented in table 3. It is noteworthy that both h and h' have an acceptable fit according to all four criteria. Hence, both models are empirically adequate (i.e., they are consistent with the data, as shown by their fit indices). CFI/TLI might be closest to 1 and RMSEA/SRMR closest to 0 in the case of h', but h has a good fit nonetheless. There are *a priori* no reasons to reject h in light of the results. Hypotheses h and h', as explanantia, are not equivalent: They provide different answers to the question regarding why there is variation in subject reports.[13] Nonetheless, they are epistemically equivalent in the sense that nothing enables us to show the superiority of one over the other as an explanans.

In both examples involving EFA and CFA, the data provide support for both hypotheses. Hence, both beliefs in h and h' are justified. However, the measurement models are mutually exclusive hypotheses (i.e., h implies $\neg h'$). Assuming that beliefs are closed under consequence, a belief in h implies a belief in $\neg h'$. Assuming the transitivity of justification principle, it follows that if a belief in h is justified, then a belief in $\neg h'$ is also justified. Consequently, minimal epistemic realism entails that researchers are justified to believe that h' is both true and false, which is absurd. Accordingly,

---

[12]See also Browne and Cudeck (1992), Hu and Bentler (1998), Bentler (2007) and Ullman (2013).

[13]Although both h and h' involve a three-factor model, the factors in these models are defined differently. In h, they are defined via four items, whereas they are defined with two items in h'. The mathematical definitions are distinct.

researchers need to deny the necessity of minimal epistemic realism. Put differently, minimal epistemic realism is not a necessary minimal epistemic standpoint insofar as it is inconsistent when there is underdetermination, and there is underdetermination, even in cases of well-confirmed and theoretically sound models.

In short, one should not assume that the data provide a justification of the belief in the model's truth, for this would also justify a belief in the model's falsity given that competing models, which may be equally well supported by the data, are mutually exclusive. This is not a problem for empiricism seeing that if h is empirically adequate, it does not imply that h' is not empirically adequate. A belief in the empirical adequacy of two mutually exclusive hypotheses or explanations can be justified via empirical data. Choosing to model reflectively is not arbitrary if researchers think that data can provide support in favor of the belief in the empirical adequacy of their hypotheses.

This example brings to light that empirical adequacy is not sufficient to justify the choice of one hypothesis over the other in structural equation modeling. In our case, there is underdetermination precisely because both hypotheses are empirically adequate, which is established by their fit indices. Hence, one might wonder how to choose between two empirically adequate hypotheses (models, explanations). As it happens, there is no universal rule allowing to choose between competing models, and everything depends on the context. For instance, if one is in the process of building the scale, then one will prefer the shorter scale for methodological reasons (e.g., one needs to minimize the time required by participants to fill out questionnaire). There are also theoretical issues to consider. For instance, it might make sense conceptually to have a cross-loading between two factors, in which case the researcher might want to rephrase the item in order to avoid any ambiguity and tap into only one factor. Similarly, it might be that the cross-loading is an idiosyncratic property of the sample, in which case it may be ignored. Besides, there are statistical elements to consider. For example, it is recommended to have at least three items per factor, so one might want to prefer h instead of h'. Further, one might want to take into consideration the balance between complexity and goodness-of-fit, as measured by (among others) the Akaike Information Criterion (AIC; Akaike 1987). Another element one might take into account is whether the model has been cross-validated (which is the case in our example). If it has, then, unless one has power issues, one should keep the entire scale (here, h would be retained). This is not meant to be an exhaustive list, but is rather meant to exemplify that

the choice between competing empirically adequate hypotheses depends on further (pragmatic) criteria that may vary across contexts (see also Peterson 2018). This issue is inherent to model selection and even realists need to consider these criteria when choosing between competing models.

4.2 Predicting abilities

Following Hood, our objective is not to determine what the epistemic attitude of researchers actually is. Rather, the objective is to determine which epistemic attitude is *necessary* to make sense of the non-arbitrariness of psychological research. Hood's argument in favor of the point that some methodological decisions do not require ontological commitments is that postulating the existence of a latent ability does not require a commitment to said ability. As it stands, this argument can be applied to Hood's second argument regarding the prediction of abilities: When trying to estimate latent abilities in participants, researchers do not need to be committed to the *existence* of these abilities. Rather, they only need to be committed to the *possibility* that these abilities exist (cf. Hood 2013, p.754). However, being committed to the possibility that latent abilities exist does not necessarily imply a commitment to the existence of said abilities. Hence, in contrast to Hood's and Borsboom's et al. positions, trying to predict latent abilities in participants is devoid of ontological commitments.

Latent abilities may or may not exist. If they do exist, then they might be identified through specific experiments and be predicted by specific tests. On the assumption that latent abilities might exist, it is intelligible that researchers try to estimate these abilities, although they might remain agnostic with respect to whether or not these abilities are actually real. Trying to predict abilities is not an arbitrary choice as long as researchers think that it is possible to justify the belief in the empirical adequacy of the hypothesis specifying the existence of latent abilities.

4.3 Longitudinal studies

A good example of a longitudinal study that can be used to model within-subject variation is a diary study. In a diary study, researchers will administrate a short questionnaire repeatedly (e.g., each day of a week, the first day of each week during one month, each hour within a day, etc.) to participants in order to evaluate the within-subject variation on the items (cf.

Reis and Gable 2000). One way to model this type of data is to perform a multilevel analysis (cf. Byrne 2012). When performing a multilevel analysis, one specifies a cluster variable that is used to regroup different measurement time points and associate them with one entity (e.g., a group, an individual). Then, one needs to specify a model representing the within-cluster variation as well as a model for the between-clusters variation. Using this type of analysis, one can model what is happening within a cluster (or individual), as well as what happens between clusters.

Underdetermination can also be used to argue against the necessity of minimal epistemic realism with respect to the choice of performing longitudinal studies that aim to model within-subject variation. As it was the case in the example presented earlier (cf. section 4.1), it is always possible to define a scenario where the data provide support for competing hypotheses that are mutually exclusive and theoretically sound. Accordingly, the argument provided in section 4.1 also applies to longitudinal studies. The only requirement needed to justify that the choice of longitudinal studies is not arbitrary is that it will yield data that provide support in favor of the belief that the hypotheses stipulating the existence of latent attributes are empirically adequate. Again, there is underdetermination and there will be mutually exclusive models that will be supported by the data. Therefore, researchers need to deny the necessity of minimal epistemic realism, otherwise both beliefs that a model is true and false would be justified. To make sense of their choice to perform a longitudinal study meant to model within-subject variation, researchers only need to be committed to methodological empiricism.

4.4 Underdetermination

It should be emphasized that we are not arguing that minimal epistemic realism is never warranted. After all, methodological realism only states that *some* methodological decisions require an appeal to minimal epistemic realism to be justified. Rather, we argue that minimal epistemic realism is not necessary to make sense of specific methodological choices in social science research. We argue against Hood's claim that minimal epistemic realism is *necessary* to make sense of the non-arbitrariness of i) choosing to model reflectively, ii) trying to predict abilities, and iii) choosing to use longitudinal research designs that model within-subject variation. Our analysis relies on the following underdetermination argument for factor analysis and, more

generally, structural equation modeling.

Underdetermination argument   In cases of underdetermination, being committed to the idea that data provide a justification of the belief in the hypothesis' truth leads to an inconsistent epistemic standpoint.

Minimal epistemic realism is neither a *minimal* or a *necessary* epistemic standpoint with respect to factor analysis given the omnipresence of underdetermination. Being committed to the idea that data provide support in favor of the belief in a hypothesis' truth would imply believing that the hypothesis is both true and false insofar as there are often (if not always) alternative competing hypotheses that are equally well supported by the data. Thus, minimal epistemic realism is not a *necessary* epistemic standpoint regarding factor analysis seeing that it is inconsistent when there is underdetermination, and there is underdetermination, even with well-confirmed and theoretically sound models. The minimal epistemic attitude researchers need to have must be able to cope with these cases. Given that minimal epistemic empiricism covers underdetermination, that minimal epistemic realism does not, and that minimal epistemic empiricism necessarily follows from minimal epistemic realism, it follows that only the former deserves to be qualified as *minimal.*

Although researchers postulate hypotheses regarding observable and latent traits, they are really dealing with models involving latent variables and subject reports. Scientific practice does not require to postulate any special relationship between latent traits and latent variables. The only minimal epistemic standpoint that is necessary to make sense of the aforementioned methodological decisions is the idea that acceptance of a theory or model requires believing in its empirical adequacy.

Incidentally, it is noteworthy that the underdetermination argument implies that *propositional justification* (see section 3.1) is actually not a suitable type of justification for minimal epistemic realism. Indeed, the empirical data and the methodology are not sufficient to guarantee that latent constructs exist. Again, the fact that a hypothesis is confirmed does not imply it is true: Confirmation of a hypothesis is not sufficient to guarantee its truth. Consequently, minimal epistemic realism rather needs to be understood in terms of weak doxastic justification, and this somewhat weaken the claim that minimal epistemic realism allows to ensure that human and social sciences are objective and rational (non-arbitrary) enterprises.

In contrast, propositional justification is perfectly suitable for minimal epistemic empiricism. Indeed, the confirmation of a hypothesis does imply that it is empirically adequate and, as such, it provides sufficient reasons to believe in the empirical adequacy of said hypothesis.

4.5   Hood's objection

Concentrating on the first argument, Hood would in all likelihood argue that empirical adequacy does not suffice to justify the choice of a reflective measurement model. After all, if Jensen and Weng aim was only

> empirical adequacy and not ascertaining the structure of mental ability, the selection of reflective measurement models would make little sense, as there are available formative models that also capture the data (Hood 2013, p.750).

The key notion underlying this argument is *causality*: In Hood's view, empirical adequacy does not suffice given that researchers such as Jensen and Weng aim to ascertain the causal structure of mental ability. As such, the choice to model reflectively is motivated by the fact that researchers aim to provide *causal* explanations.

Accordingly, the real challenge seems to determine how one could motivate the choice of a reflective measurement model over a formative model if the aim is not to uncover the true causal structure underlying the psychological phenomenon under study but is rather to be solely 'empirically adequate'. That is, a causal hypothesis is a hypothesis about the real structure of the world, so if one wants to *explain* the between-subjects variation on items through a latent variable, then one needs to assume that it is possible to justify this causal hypothesis, hence minimal epistemic realism needs to be assumed. The only way to rationally justify the choice of a reflective measurement model would be to appeal to the causal structure of psychological phenomena. The reflective model is chosen because the factors are taken as causal explanation of the between-subjects variation on the items (cf. Hood 2013, p.749).

It is noteworthy that this idea, namely that causal explanation is related to the physical structure of the world, is implicit within the philosophy of science literature on psychology (cf. Borsboom et al. 2003, pp.208-9, 2004, pp.1061-4; Michell 2005, p.289; Hood 2013, p.749).[14] On this subject, Wright

---

[14]See also Bechtel and Abrahamsen (2005) as well as Craver and Bechtel (2007).

writes that if one wants to avoid realism in psychology, one needs to address the presupposition that

> the causality of ordinary psychological explanations requires their hypotheses to carry a content that already puts them out of the running for minimalist construal: that there is an objective and at bottom wholly physical causal order in the world and that once a discourse ventures causal claims, it must be entered into the competition for correct depiction of aspects of this causal order and sink or swim accordingly (Wright 2002, p.220).

Therefore, the biggest challenge one faces when trying to untangle the epistemic assumptions underlying psychological research from an empiricist perspective seems to be to provide an account of explanation without appealing to minimal epistemic realism or causality.

Let us start from minimal epistemic empiricism. Hood would ask: How could one rationally justify the choice of a reflective measurement model? The answer to that question is that the choice of a reflective measurement model is justified via the rationale behind scientific inference and hypothesis testing: Factors are assumed as explanations of the between-subjects variation on the items. The researcher, however, only needs to assume that it is possible to justify the belief in the empirical adequacy of the model. There is no need to assume that it is possible to justify the belief in its truth.

Hood would then reply: By assuming that factors can explain the between-subjects variation on the items, researchers are making causal claims and, thus, they at least need to be committed to the idea that it is possible to justify (beliefs in) these claims. Empirical adequacy does not suffice because they are aiming at causal structures. Causal structures are inherent properties of reality. If researchers postulate a causal structure, they need to assume minimally that the causal structure is the appropriate one. In other words, they need to assume that it corresponds to reality or, similarly, that the claim about the causal structure is true. Hence, empirical adequacy does not suffice to justify the choice of the reflective measurement model, researchers need truth.

4.6   Answering Hood's objection

Explanations in science are generally understood as causal explanations (e.g. Woodward 2003; Strevens 2008). Given the choice between a causal and

a non-causal explanation, Hood might argue that the fact that researchers elected for the causal explanation (that involves mathematics) indicates a commitment to minimal epistemic realism. But this choice is idiosyncratic and is not necessary to make sense of the choice to model reflectively.[15] Without going as far as to argue in favor of a non-causal account of the explanatory power of factor analysis (cf. Lange 2013, Skow 2014), it can be argued that postulating that researchers are trying to establish claims regarding the causal structure of psychological phenomenon is an unnecessary assumption.

Recall the distinction between observable/latent traits and subject reports/latent constructs. Only traits can have causes or effects. When appealing to latent traits to provide causal explanations, one is not in the realm of factor analysis. Reflective measurement models are meant to provide support in favor of the belief (in the empirical adequacy of the hypothesis) that latent traits can causally explain empirical phenomena. Appealing to causal explanations is unnecessary with respect to the choice to model reflectively. As it stands, one only needs a minimal (pragmatic) account of explanation. Scientific theories in social sciences have pragmatic and empirical aspects. As such, an explanation can simply be understood as an empirically adequate answer to a why question, and this answer can as well involve mathematics (cf. Baker 2005). It does not require any claim regarding what structure in the world brings about the phenomenon. Accepting a measurement model as an explanation that involves mathematics only requires a belief in its empirical adequacy.

As we mentioned earlier, empirical adequacy does not suffice to motivate the choice between two competing empirically adequate explanations in structural equation modeling. There are external elements to consider (e.g., theoretical, statistical, methodological, etc.), and the pragmatic aspect of these elements will be evaluated differently depending on the situation. As van Fraassen (1980, p.13) argued, realists tend to attribute an objective validity to explanations, an objective validity that anti-realists simply cannot grant. But realists should agree that explanations are empirically adequate answers to why questions, and that there are pragmatic elements to consider when evaluating competing explanations. Only this conception is necessary

---

[15]An operationalist, for instance, might accept the notion of a distinctively mathematical explanation (cf. Lange 2013), in which case the explanation of the between-subjects variation on the items, which is an empirical phenomenon, would rely solely on the statistical model provided by the analysis.

to make sense of the choice to model reflectively.

Besides, even if one assumes that latent traits can causally explain empirical phenomena, this does not imply a commitment to minimal epistemic realism. Even under this assumption, it follows from the underdetermination argument that assuming minimal epistemic realism will lead to an inconsistent epistemic standpoint. As such, minimal epistemic realism does not hold in this context and, therefore, it does not follow from endorsing a causal account of explanation (at the level of traits).

5   Closing remarks

To conclude, we provided an empiricist account of methodological choices in social science research, thus answering the challenge raised by philosophers such as Hood, Leplin and Boorsboom et al. We argued against Hood's claim that minimal epistemic realism is necessary to rationally justify the choices to i) model reflectively, ii) trying to predict abilities, and iii) perform longitudinal research designs meant to model within-subject variation. We argued that the issue of model selection and, incidentally, underdetermination in social science research, give us grounds to reject Hood's thesis. Following van Fraassen, we defended the thesis of methodological empiricism and argued in favor of the necessity of minimal epistemic empiricism to make sense of the non-arbitrariness of methodological choices in social sciences that use structural equation modeling: Acceptance of a model only requires, as belief, that it is empirically adequate.

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52(3):317–332.

Baker, A. (2005). Are there genuine mathematical explanations of physical phenomena? *Mind*, 114(454):223–238.

Ball, B. (2012). Knowledge is normal belief. *Analysis*, 73(1):69–76.

Bechtel, W. and Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2):421–441.

Bentler, P. M. (2007). On tests and indices for evaluating structural models. *Personality and Individual Differences*, 42(5):825–829.

Bergmann, G. and Spence, K. W. (1941). Operationism and theory in psychology. *Psychological Review*, 48(1):1–14.

Bhakthavatsalam, S. and Cartwright, N. (2017). What's so special about empirical adequacy? *European Journal for Philosophy of Science*, 7(3):445–465.

Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge University Press.

Borsboom, D. (2008). Latent variable theory. *Measurement*, 6(1-2):25–53.

Borsboom, D. and Mellenbergh, G. J. (2004). Why psychometrics is not pathological: A comment on Michell. *Theory & Psychology*, 14(1):105–120.

Borsboom, D., Mellenbergh, G. J., and van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110(2):203–219.

Borsboom, D., Mellenbergh, G. J., and van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111(4):1061–1071.

Browne, M. W. and Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21(2):230–258.

Byrne, B. M. (2012). *Structural Equation Modeling with M*plus. Routledge.

Chellas, B. F. (1980). *Modal logic: An introduction*. Cambridge University Press.

Craver, C. F. and Bechtel, W. (2007). Top-down causation without top-down causes. *Biology & Philosophy*, 22(4):547–563.

Cronbach, L. J. and Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4):281–302.

Deci, E. L. and Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian psychology/Psychologie canadienne*, 49(3):182–185.

29

Edwards, J. R. and Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, 5(2):155–174.

Feldman, R. (2004). Foundational beliefs and empirical possibilities. *Philosophical Issues*, 14(1):132–148.

Field, A. (2009). *Discovering statistics using SPSS*. Sage, 3rd edition.

Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences*, 3(1):63–73.

Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23(6):121–123.

Glymour, C. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. MIT Press.

Halpern, J. Y., Samet, D., and Segev, E. (2009). Defining knowledge in terms of belief: The modal logic perspective. *The Review of Symbolic Logic*, 2(3):469–487.

Hempel, C. G. (1943). A purely syntactical definition of confirmation. *The Journal of Symbolic Logic*, 8(4):122–143.

Hood, S. B. (2008). Comments on Borsboom's typology of measurement theoretic variables and Michell's assessment of psychometrics as "pathological science". *Measurement*, 6(1-2):93–97.

Hood, S. B. (2013). Psychological measurement and methodological realism. *Erkenntnis*, 78(4):739–761.

Horwich, P. (1996). Realism and truth. *Noûs*, 30(10):187–197.

Hu, L.-t. and Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4):424–453.

Israel, H. and Goldstein, B. (1944). Operationism in psychology. *Psychological Review*, 51(3):177–188.

Jackson, F. and Pettit, P. (1990). In defence of folk psychology. *Philosophical Studies*, 59(1):31–54.

Jensen, A. R. and Weng, L.-J. (1994). What is a good g? *Intelligence*, 18(3):231–258.

John, O. P. and Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of Personality: Theory and Research*, pages 102–138. Elsevier, 2nd edition.

Johnson, K. (2016). Realism and uncertainty of unobservable common causes in factor analysis. *Noûs*, 50(2):329–355.

Kline, R. B. (2005). *Structural Equation Modeling*. The Guilford Press.

Kline, R. B. (2013). Exploratory and confirmatory factor analysis. In Petscher, Y. and Schatsschneider, C., editors, *Applied Quantitative Analysis in the Social Sciences*, pages 171–207. Routledge.

Lange, M. (2013). What makes a scientific explanation distinctively mathematical? *The British Journal for the Philosophy of Science*, 64(3):485–511.

Leplin, J. (1986). Methodological realism and scientific rationality. *Philosophy of Science*, 53(1):31–51.

Leplin, J. (1987). Surrealism. *Mind*, 96(384):519–524.

Leplin, J. (1997). *A Novel Defense of Scientific Realism*. Oxford University Press.

Mellenbergh, G. J. (1994). Generalized linear item response theory. *Psychological Bulletin*, 115(2):300–307.

Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology*, 88(3):355–383.

Michell, J. (2003). The quantitative imperative positivism, naive realism and the place of qualitative methods in psychology. *Theory & Psychology*, 13(1):5–31.

Michell, J. (2004). The place of qualitative research in psychology. *Qualitative Research in Psychology*, 1(4):307–319.

Michell, J. (2005). The logic of measurement: A realist overview. *Measurement*, 38(4):285–294.

Michell, J. (2013). Constructs, inferences, and mental measurement. *New Ideas in Psychology*, 31(1):13–21.

Muthén, B. and Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological methods*, 17(3):313–335.

Peterson, C. (2018). Accommodation, prediction and replication: Model selection in scale construction. *Synthese*. https://doi.org/10.1007/s11229-017-1660-0.

Psillos, S. (1999). *Scientific realism: How science tracks truth*. Routledge.

Reis, H. T. and Gable, S. L. (2000). Event-sampling and other methods for studying everyday experience. In Reis, H. T. and Judd, C. M., editors, *Handbook of Research Methods in Social and Personality Psychology*, pages 190–222. Cambridge University Press.

Skow, B. (2014). Are there non-causal explanations (of particular events)? *The British Journal for the Philosophy of Science*, 65(3):445–467.

Spirtes, P., Glymour, C., and Scheines, R. (1991). From probability to causality. *Philosophical Studies*, 64(1):1–36.

Spirtes, P., Glymour, C., and Scheines, R. (2000). *Causation, Prediction, and Search*. MIT Press.

Stevens, S. S. (1939). Psychology and the science of science. *Psychological Bulletin*, 36(4):221–263.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103(2684):677–680.

Strevens, M. (2008). *Depth: An account of scientific explanation*. Harvard University Press.

Thurstone, L. L. (1947). *Multiple factor analysis*. University of Chicago Press.

Thurstone, L. L. (1954). An analytical method for simple structure. *Psychometrika*, 19(3):173–182.

Trépanier, S.-G., Fernet, C., and Austin, S. (2013). Workplace bullying and psychological health at work: The mediating role of the satisfaction of needs for autonomy, competence and relatedness. *Work & Stress*, 27(2):123–140.

Turri, J. (2010). On the relationship between propositional and doxastic justification. *Philosophy and Phenomenological Research*, 80(2):312–326.

Ullman, J. B. (2013). Structural equation modeling. In Tabachnick, B. G. and Fidell, L. S., editors, *Using Multivariate Statistics*. Pearson, 6th edition.

Van den Broeck, A., Vansteenkiste, M., De Witte, H., Soenens, B., and Lens, W. (2010). Capturing autonomy, competence and relatedness at work: Construction and initial validation of the work-related basic need satisfaction scale. *Journal of Occupational and Organizational Psychology*, 83(4):981–1002.

van Fraassen, B. C. (1980). *The Scientific Image*. Clarendon Press.

Velicer, W. F. and Jackson, D. N. (1990). Component analysis versus common factor analysis: Some further observations. *Multivariate Behavioral Research*, 25(1):97–114.

Waters, R. H. and Pennington, L. A. (1938). Operationism in psychology. *Psychological Review*, 45(5):414–423.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91(3):461–481.

Woodward, J. (2003). *Making things happen: A theory of causal explanation.* Oxford University Press.

Wright, C. (2002). What could antirealism about ordinary psychology possibly be? *Philosophical Review*, 111(2):205–233.