A Dispositionalist Solution to the “Kripkenstein” Rule-Following Paradox

鄭凱元
中正大學哲學系

I. What is the “Kripkenstein” rule-following paradox?

Kripke (1982) constructs from Wittgenstein’s *Philosophical Investigation* (“KW” in short, or “Kripkenstein”) a skeptical argument which leads to the skeptical paradox that there is no fact of the matter about an individual speaker that can determine whether he follows one rather than another rule. KW thus concludes that there is no such thing as rule-following (or meaning).

The skeptical argument proceeds by elimination: several promising facts that may constitute rule-following are considered and rejected one by one. Among the facts that are examined are a person’s past usage, employment of a set of instruction, disposition of a certain kind, qualia, etc. None is found appropriate to constitute rule-following. Hence, it must be the case that, KW urges, there is no such thing as rule-following to begin with.

The skeptical conclusion is extremely counterintuitive and implausible. Many Philosophers have tried to argued against KW’s skeptical paradox for the last two decades. No consensus, however, has yet been reached which position or approach is likely to succeed in responding to KW’s skeptical paradox. In this paper, I will offer a dispositionalist solution to the paradox in question. The aim is to vindicate the reality of rule-following and meaning.

I will assume that readers are already familiar with the rule-following paradox, and hence will not explain KW’s skeptical argument in detail. Those who wish to know more about the skeptical argument can consult Kripke’s 1982 book. This paper can nonetheless be read without having much prior knowledge about KW’s sceptical argument.

This paper is a draft based on some parts of my dissertation. Please do not cite or circulate. Comments and feedback are most welcome!
II. How does KW argue against the dispositional account of rule-following?

Three major objections:
1. The error problem (often known as the normativity problem)
2. The finitude problem
3. The first-person epistemic problem

1. What is the error problem?

To explain the error problem, we have to first explain the counterfactual idea of a disposition held by KW. Consider the following two statements:
(f) The piece of glass is fragile.
(g) If the piece of glass on the table were to fall on the floor, it would break.
The statement (f) ascribes a disposition, i.e., fragility, to a piece of glass.
The statement (g) is a counterfactual conditional that consists of two indicative statements: the antecedent describes a certain testing condition, while the consequent describes a certain reaction of the piece of glass when put in the testing condition. If (g) can be viewed as a necessary and sufficient condition for (f), then (g) offers a reductive explanation of (f). The motivation underlying the counterfactual analysis of a disposition is to give a reductive explanation of something mysterious (e.g., a disposition) by something that is less mysterious (e.g., two observable events).

Similarly, meaning (or rule-following) is first equated with a semantic disposition, and then statements about the disposition is further analyzed in terms of a set of counterfactual conditions:
(m) S meant plus by “+”.
(d) S had a certain disposition with respect to the operation of the sign “+”.
(c) If S had been queried about “m + n”, he would have responded with the sum, where the triple of (m, n, r) is an instance dictated by the plus rule.

In order for the counterfactual (c) to analyze the disposition ascription (d) and hence reductively explain the meaning ascription (m), the infinite set of instances of (c) has to be necessary and sufficient for (m).

This, however, turns out to be very difficult to accomplish. For a person might follow the plus rule, but fails to answer “125” to “68 + 57”, for he might be
drunk (in this case, (c) is not necessary for (m)). A person might follow a non-standard rule, such as the quus rule, but wrongly answers, say, “125” to “68 + 57”, for he might be too tired (in this case, (c) is not sufficient for (m)). In both cases, we would say that the person’s failing to give “correct” answers is due to the potential occurrences of intervening factors, such as drinking too much alcohol, or being too tired, etc.

The challenge of the error problem can be viewed as the challenge of offering a necessary and sufficient condition for (c) that is indicative of a rule-following disposition.

The typical way to meet this challenge is to supplement the antecedent of a counterfactual with a normality condition from which (together with the antecedent) the consequent of the counterfactual can be derived. A satisfactory normality condition is one in which all the potential intervening factors are listed and excluded. This proves hard to be done, because of the problem of vacuity or circularity. Consider (c1) and (c2) as follows:

(c1) If S had not been drunk, not tired, not received a heavy blow on the head, and not in any other unfavorable conditions, S would have responded with the sum “r” when queried “m + n”.

The conditional (c) is obviously vacuous, for what “any other unfavorable situations” are is not clearly specified. Due to this difficulty, we might be tempted to give the following formulation of a counterfactual:

(c2) If nothing had interfered, S would have responded with the sum “r” when queried “m + n”.

(c2) is clearly trivially true, for the insertion of “had nothing interfered” simply amounts to presupposing that the original counterfactual (c) is true, and hence presuppose the truth of the meaning-ascribing sentence (m).

The error problem for a dispositional account of rule-following is thus fundamentally a task of saving a counterfactual conditional from vacuity or circularity, given the counterfactual analysis of a disposition held by KW’s sceptic.
2. What is the finitude problem?

The finitude problem, as KW conceives it, is a problem arising from the fact that a disposition is finite. KW points out that a person’s rule-following disposition is capable of producing only finitely many calculations. However, a rule or meaning dictates infinitely many calculations. How then can a finite disposition encompass the infinitely many judgments that are sanctioned by a rule? The consequence seems to be that a disposition and following a rule cannot be equated.

Put in a different way, my disposition, being manifestable only finitely many times in the past due to my finite capacities, is subsumed under numerous rules which are distinct and mutually incompatible, and whose instances diverge right after the point of my finite performances. How then can I (or we) be justified in claiming that one particular rule rather than another of these numerous ones has been followed during my life-time? There appears to be an under-determination problem: one’s finite disposition to produce behaviors under-determine the rule we follow. Underlying the finitude problem is, hence, the under-determination problem, which faces the dispositional account of rule-following.

3. What is the first-person epistemic problem?

The first-person epistemic problem is derived from the intuitive idea that the meaning or rule we grasp guides and justifies our application of a term or rule-following behaviors. A machine might be able to add, by behaving in a way that exhibits a certain rule-conforming pattern. What distinguishes the machine and a competent adder like me is that when I add, I seem to be directly guided by the rule that I follow, and I am capable of citing the rule that I consciously grasp to justify my performances. A disposition, on the other hand, is something whose identity condition can only be known indirectly through inference from the behaviors manifested under various conditions. How then can a disposition be identified as the fact that constitutes meaning or rule-following? The challenge for a dispositional account of rule-following is to explain how one could have conscious contact or direct access to one’s own underlying disposition to rule-following.

The first-person epistemic problem is not widely noted. But it is, in my view, a very important problem that any plausible dispositional account has to address. I deal with this problem in Chapter VIII of my thesis. I hope to present it in
another occasion.

III. Previous dispositionalist attempts to solving the error problem

1. Shogenji’s appeals to the concept of stable rule-following behaviors to serve as the criterion of correct behaviors in rule-following:

In Shogenji’s account, it is one’s disposition to produce stable performances that determine the rule being followed. This criterion does not appear to be circular at the first glance, because what those stable behaviors a person is disposed to produce are left entirely open. For example, if a subject stably responds “1” to “4 + 3”, then the (basic routine) rule he follows is \(4 + 3 = 1\) according to Shogenji’s account.

Upon reflection, however, epistemic doubt arises as to why a particular set of performances (i.e., stable ones) is favored over another set of performances (e.g., unstable ones). Reason has to be provided to justify the choice. Otherwise, the account is bound to be presuppose which rule is followed. Moreover, even if a reason is given, further question still arises why the provided reason stands. An infinite regress for justification would thus occur.

2. Yallowitz’s appeal to one’s verdicts about one’s own rule-following behaviors to serve as a criterion of correct and incorrect behaviors:

A behavior is correct if the person who produces it says that it is correct, unless there is evidence to the contrary. In Yalowitz’s view, KW is wrong in bracketing the presumption of the first-person authority, for the presumption of first person authority is as fundamental to our commonsense picture as normativity (or the correctness conditions of a term’s usage).

The problem with this approach, however, is that one’s verdicts, either in the form of a judgment or a feeling of certainty, can guide one right but can also guide one wrong when following a rule. It cannot be decisively determined which verdict is right or when a verdict is right.

It is precisely because we cannot be sure which rule-following behavior is correct and hence which rule is followed that KW’s sceptic suggests that we give up first-person authority as a means o fix the rule being followed. Yalowtiz’s emphasis on the commonsense importance of first-person authority of one’s disposition to make judgments about one’s own behavior simply leaves the error
Consequently, a standard for distinguishing the correct behaviors from incorrect ones remains needed to fix the rule being followed.

3. Martin & Heil’s taking a metaphysical approach to solving the error problem:

Martin & Heil does not confront the error problem directly by offering a plausible set of counterfactuals. Instead, it solves the problem by discarding KW’s counterfactual analysis of dispositions. In Martin & Heil’s view, the error problem is an unfortunate consequence of KW’s holding a mistaken counterfactual conception of dispositions. Once we have a correct understanding of the ontological character and status of a disposition, it can be shown how a disposition is an appropriate fact that constitutes rule-following.

In KW’s counterfactual construal of dispositions, rule-following is ontologically nothing by a set of infinite potential behaviors exhibited under an infinity of possible testing conditions. In this view, a disposition is a mere *possibilia*, while its manifestations actual and real. Martin & Heil reverse the order of explanation by arguing that a disposition is actual, while its manifestations are mere *possibilia*. In Martin & Heil’s realist view, a disposition is an actual property of the object (or person) that possesses it.

Martin (1994) offers an important argument against the conditional analysis of a disposition. He asks us to consider first the following disposition ascription A:

A: The wire is live.

And then, consider the conditional B:

B: If the wire is touched by a conductor then electric current flows from the wire to the conductor.

On the reductionist view, B is meant to be A’s *analysans*. It turns out, however, that B is neither necessary nor sufficient for A. Martin shows this by considering an electro-fink machine that is connected to the wire. The function of the electro-fink machine is to detect when the wire is about to be touched by a conductor and can instantaneously make a live wire dead or a dead wire live. An electro-fink thus produces the following two cases. In the first case, imagine that the wire is dead and not touched by a conductor. In that case, A is false but B is true. B is true because the wire’s connecting to an electro-fink machine ensures that if the wire were to be touched by a conductor then electric
current would flow to the conductor. As a result, B is not sufficient for A. In
the second case, imagine that the wire is live and not touched by a conductor. A
is thus true while B is false. B is false because, once again, the wire is connected
to an electro-fink machine, which renders it the case that if the wire were to be
touched by a conductor, then current would not flow from the wire to the
conductor. Hence, B is not necessary for A. The electro-fink case thus shows that
no conditional is necessary or sufficient for a disposition-ascribing statement.

The result of the two cases of the wire example can be generalized as follows.
A disposition ascription is not equivalent to a conditional, because the
disposition
ascription could be true while the conditional false and the disposition ascription
could be false while the conditional true. In Martin’s view, a conditional analysis
of a disposition is “clumsy” and “inexact.” Interfering factors such as an
electro-fink machine might enter into causal contact with a disposition such as
the electric power of a wire and render the disposition unable to manifest itself.
Martin’s primary concern is hence not only to point out the inadequacy of a
conditional analysis of a disposition, but also to advance our realist intuition
about a disposition.

Martin & Heil’s metaphysical considerations are, in my view, insightful and
plausible. They capture our important intuition that even if an agent is not put
under a testing condition and hence does not manifest a behavior, the
rule-constituting disposition is inherent in the agent and is there-and-ready for a
suitable manifestation. This is similar to our intuition about, say, a bridge. Even
if no car is passing over it, the bridge’s disposition to support cars is inherent in
the bridge and is there-and-ready for a suitable manifestation. If what a
counterfactual analysis does is merely to reduce a disposition to a pair (or a set of
pairs) of testing conditions and manifestations, then the analysis is very likely to
be wrong. We can easily imagine that a bridge still possesses its enduring
capacities even though no cars ever drive over it.

Philosophers favoring a counterfactual analysis of the disposition, however, may
argue that they do not deny that a bridge has an enduring capacity even though
no cars ever drive over it. Rather, their claim is only that the disposition
ascription “the bridge has the capacity of supporting cars at t, when no cars were
driving over it” has to be analyzed in terms of the counterfactual conditional “If
any car had been driving over the bridge at t, it would have held”. To me, this
argument is similar to Berkeley’s treatment of the existence of an unseen object. Berkeley does not deny that objects could exist unseen; he merely contends that the existence statement “An object O exists unseen at t” must be treated as being equivalent to the counterfactual conditional “If someone had been looking at t, he would have had O-ish experiences.” The consequence of his treatment is to reduce the existence of an object to a counterfactual conditional, which amounts to denying that an object can exist independently of a potential perceiver who is to perceive the object under certain conditions. As Berkeley himself writes: “…the absolute existence of unthinking things without any relation to their being perceived, that seems perfectly unintelligible. Their \textit{esse is percipi}, nor is it possible they should have any existence out of the minds or thinking things which perceive them.”

Berkeley’s counterfactual analysis of the existence of objects is, however, counter-intuitive. Although we may agree that the existence statement P: “There is a table in my room” has some logical relation to the counterfactual statement P’: “If someone entered my room, he would or could have the experience of seeing a table,” our commonsense intuition is that the second statement P’ is the outcome of the first statement P. That is, we intuitively think that the existence of the table does not require its being perceived by an agent. Rather, the existence of the table is directly responsible for an agent’s being able to perceive it under certain conditions, and the existence of the table also explains the truth of the counterfactual statement P’. Berkeley’s semantic analysis of existence statements may be onto something, but the ontological consequences he draws from his analysis are misguided. Similarly, our realist intuition about a disposition should be suitably appreciated and kept salient. Contrary to counterfactual reductionists’ claim, my claim is, therefore, that ascriptions of dispositions are categorical, that they refer to real properties of objects which possess them, rather than counterfactual.

Rejecting a counterfactual construal of a disposition based on realist insights still leaves us with a puzzle: In what sense is a disposition real and actual, as Martin & Heil claim? If a disposition is not to be equated with a counterfactual conditional, then what can it be identified with? Is a disposition to be identified with the causal basis of an object or subject that possesses the disposition? Or does a disposition have a causal basis at all? If a disposition has a causal basis but is distinct from its causal basis, then what are their relations, and how do they interact? If a real disposition is not identical with its causal basis, then the
disposition will not be causally efficacious. In that case, in what sense can it be real? These are difficult and controversial questions. I don’t think they have to be answered in the context of responding to KW’s skeptical paradox. At this stage I will stick to a minimalistic realist view about a disposition. Namely, a disposition is real in the sense that a disposition ascription is not to be reduced to or identified as a mere counterfactual conditional that describes manifestations under various conditions. A disposition is something that makes a counterfactual conditional true.

**Martin & Heil’s solution to the error problem:**

Martin & Heil’s emphatic distinction between a disposition and its manifestations has an important implication for their solution to the error problem. To follow a rule, as noted, is to possess a disposition of a certain sort, which is in turn characterized as one that occupies a particular position in a holistic web of reciprocal and inhibitory dispositional partners. For example, when put under a testing condition, an agent may not manifest a particular semantic disposition because of the absence of some reciprocal dispositional partners (e.g., the agent is deprived of oxygen for, say, two minutes) or because of the presence of some inhibitory dispositional partners (e.g., the agent has earlier taken some sleeping pills). Nevertheless, there still exists a meaning-constituting fact, a real disposition occupying a certain position in the holistic structure of dispositions.

In Martin & Heil’s metaphysical picture, therefore, errors occur because a particular disposition fails to manifest itself due to blocking factors, so are not at all critical or even relevant for the determination of the rule that is being followed. What is critical for rule-determination is instead the determination of the underlying disposition that is embodied by an agent. The disposition should not be equated with its successful or unsuccessful manifestations, or omissions. The fundamental mistake, according to Martin & Heil, of a counterfactual analysis is that it confuses disposition with manifestations. As a result, the theorist wrongly looks for an epistemic criterion for classifying manifested behaviors into correct and incorrect, and thus faces the thorny error problem. The

---

1 Major positions and discussions of the metaphysics of dispositions can be found in the following two excellent books: *Dispositions: A debate*, by Tim Crane (ed), 1996, and *Dispositions*, by Stephen Mumford, 1998.
right way to determine the rule being followed, according to Martin & Heil, is to
directly appeal to the web structure of a particular real disposition under scrutiny. The error problem re-surfaces in Martin & Heil's realist dispositional account:
The difficulty for Martin & Heil's dispositional account, however, is that the
error problem does not go away. Martin & Heil solve the error problem by
arguing that an error occurs because some intervening factors prevent a particular
disposition from manifesting itself. So, in order to identify an erroneous act, we
have to know whether the act is performed absent blocking factors. How then are
we to identify such a condition without having to presuppose what to count as
blocking factors, and hence presuppose which real rule-constituting disposition is
embodied?
Let me elaborate the difficulty a little more. The difficulty that Martin & Heil’s
view encounters is an epistemological one. Martin & Heil emphasize the
distinction between an underlying disposition and its manifestations under
certain conditions. However, they say relatively little about how we know what
disposition an agent possesses, if he has one. We cannot point to a disposition
and see it directly, because it is theoretical and not directly visible (on their
account). So, if there is any way we can know what a disposition is like, there
has to be some intimate connection between a disposition and its manifestations
so that we may know unambiguously what disposition it is via the mediation of
its observed manifestations under various conditions. But then the error problem
emerges again. For only “suitably” (or correctly) manifested behaviors can tell us
which disposition is embodied by an agent, where “suitably” means that the real
disposition is not being blocked by the presence of any inhibitory dispositional
partner or by the absence of reciprocal dispositional partner. The skeptic can
reasonably doubt that we can decide without circularity whether a blocking
factor is present. Because of the interlocking and interdependent relations
between a particular real disposition and its dispositional partners, it does not
seem plausible that one can identify a blocking factor without presupposing what
a particular disposition is in place. As a result, the rule that is being followed
cannot be determined without circularity.

IV. A Solution to the Error Problem in Martin & Heil’s Realist Dispositional
Account of Rule Following
My claim is that the error problem that is left unresolved in Martin & Heil’s
account can be solved by adding the following two supplementary points. The
first point is that ascribing a real disposition to an object or subject is equivalent
to an act of hypothesizing or theorizing. The second point is that such theorizing
is abundant in scientific practice, and it is perfectly legitimate to infer the underlying property of an object. I will argue for these two points respectively.

Hypothesizing about Particular Dispositions and the Resulting Problem of Vacuity of Provisoes and Ceteris Paribus Clauses

We have granted that the realist view of a disposition proposed by Martin & Heil is better than a view which reduces a disposition to a counterfactual conditional. We thus agree that a real disposition must be suitably distinguished from its manifestations. However, the underlying real disposition that is possessed by an object or subject is invisible and hence unknown. How then do we know what it is?

Hempel (1977) gives a lucid DN model of a dispositional explanation in which ascribing a particular disposition to an object or subject is shown to be a matter of hypothesizing and can figure in a scientific inference. The DN model of explanation can be schematically represented as follows:

(DN)
(C1) i was in a situation of kind S.
(C2) i has the (dispositional) property M.
(L) Any x with the property M, if it were in a situation of kind S, would behave in manner R.

(E) i behaved in manner R.

In (DN), (C1) describes an initial condition, and (C2) contains a theoretical term M that ascribes a dispositional property to i. (L) states a theoretical principle, which contains M, in the form of a conditional (x)(Mx -> (Sx -> Cx)). And finally, (E) is a conclusion stating the behavior of i. The derivation of (E) has a deductive-nomological (D-N for short) form, which allows the object i’s actual or potential behavior to be explained or predicted. Such an argument in a DN form helps us see how ascribing a disposition does its explanatory work.

In adopting Hempel’s DN model of dispositional explanation, there is a question whether stating the theoretical principle (L) as a conditional involves reducing the disposition ascription (C2). Hempel is not fully explicit on this. However, even if he holds a reductivist view, we can still, in my view, construe a disposition ascribed by (C2) realistically, and regard (L1) as a consequence of an object possessing M. What is crucial and relevant in the context of solving KW’s error problem is this: Hempel’s DN model of dispositional explanation helps us see that ascribing a disposition to an object is equivalent to forming a hypothesis
like (C2) and that a conditional like (L1) follows from an object’s having the particular real disposition that the hypothesis (C2) is about. Hempel’s characterization of a dispositional explanation of an empirical phenomenon can be similarly applied to a realist dispositional account of rule-following as follows:

(DN*)

(C1*) A subject s was presented with the question “m + n = ?”.

(C2*) s has the disposition P with regard to operating the sign “+”.

(L*) Any subject x with the disposition P, if he were presented with the question “m + n = ?”, would answer the sum r.

(E*) s answered the sum r.

As Hempel points out, however, a DN model of dispositional explanation of an empirical phenomenon faces the problem of provisos. The problem of provisos has to do with the fact that there are cases in which all the premises of a DN argument are true while the conclusion false, and thus the DN argument is invalid. For example, there might be disturbing factors present which causes the actual observed result to conflict with the theoretically derived result stated in (E). However, the theory stated in the premises (C2) and (L) does not ensure that those potential disturbing factors are absent. Hence, the inference of DN model has to presuppose as a premise that no such disturbing factors are present. Hempel calls such a premise as a “proviso”. A proviso is not explicitly sated, but is essentially employed in a DN dispositional explanation of an empirically phenomenon.

The problem of a proviso is that it is only vaguely expressed, and hence that whenever an object or subject’s behavior is observed to differ from the predicted result, one can always claim that the implicit and vague proviso is violated (i.e., that some disturbing factors are present). Unless a criterion can be given as to when a proviso is satisfied and when it is violated, a dispositional explanation will never be falsifiable because it inherently contains an unstated escape clause. This renders a hypothesis that ascribes a disposition to an object or subject impossible to confirm or disconfirm. As a result, which rule a subject follows remains undetermined. The error problem that arises in Martin & Heil can thus be characterized as the problem of provisos in Hempel’s DN model of dispositional explanation.
A proviso is similar to a normality condition typically added to a traditional counterfactual conditional that is to analyze a disposition ascription. Normality conditions are those under which no disturbing or intervening factors enter into causal relation with the particular disposition under scrutiny. Therefore, whenever an object’s or subject’s disposition manifests itself in a way which conflicts with what the counterfactual conditional describes, it can simply be claimed that conditions are not normal, and the manifested behaviors cannot be attributed to the normal functioning of the disposition.

The function of provisoes in theoretical inferences, such as those characterized by a DN model, is similar to that of *ceteris paribus* clauses that are widely employed to protect laws of special sciences from counterexamples. With a *ceteris paribus* clause attached to a law, counterexamples to the law can be claimed to occur because *other things* are not *equal*. Adding a *ceteris paribus* clause to a law is tantamount to making an assurance that when disturbing factors are absent, a nomological pattern of behaviors will be observed as predicted. It is clear then that the problem of unfalsifiability that confronts Hempel’s DN model of dispositional explanation also confronts a *ceteris paribus* law, and a counterfactual conditional supplemented with normality conditions. The difficulty in each of these cases arises from the fact that there is serious vacuity inherent in a proviso, a *ceteris paribus* clause, or a normal condition. Thus, a solution to the problem of vacuity in any one of the three cases applies to the other two as well.

**A Solution to the Problem of Vacuity of Provisoes, Ceteris Paribus Clauses, and Normality Conditions**

Pietroski & Rey\(^2\) argue for the importance of *ceteris paribus* clauses in science. In their view, *ceteris paribus* clauses are indispensable tools for the abstraction and idealization much needed in scientific theorizing. Pietroski & Rey stress the indispensable role played by cp-clauses in scientific theorizing, but they also recognize that vacuity of cp-clauses is ineliminable. They thus offer a sufficient condition for the non-vacuity of cp-laws. If their attempt is successful, then the problem of unfalsifiability inherent in scientific theories can be solved, and so can the error problem in a dispositional account of rule-following.

---

The core of Pietroski & Rey’s proposal is that, if all apparent counterexamples to a cp-law can be explained by independent factors that are actively interfering with the nomologically causal relation described by the law, then the cp-law is non-vacuous. Independent interfering factors are independent in the sense that, roughly, they can be explained by theories that invoke laws other than the one under investigation (we will go into this in more detail shortly). Pietroski & Rey express their proposal metaphorically, “cp-clauses are cheques written on the bank of independent theories.” Just as writing a cheque carries a commitment that the bank where the cheque is deposited will cash it, adding a ceteris paribus clause carries with it an explanatory commitment that other theories than the one being offered will explain counterexamples to the law. In Pietroski & Rey’s “cheque-cashing” proposal, therefore, the substance and warrant of a ceteris paribus clause are derived from the substance and warrant of those other theories. If the ceteris paribus clause of a law can be “cashed” by other theories, then the cp-law is a good one. If not, then the cp-law is genuinely vacuous and better jettisoned.

In Pietroski & Rey’s sufficient condition, it is crucial that counterexamples or exceptions to a cp-law must be explainable by citing independent interfering factors. But what exactly does the notion of independence mean here? We may take clairvoyance for example to illustrate it. Some people may believe that the power of clairvoyance exists, and accept the following cp-law as true: cp (if medium x predicts p, then p). Suppose that experiments are conducted to test whether the cp-law in question is true, and it turns out that people claimed to be clairvoyant fail to exhibit clairvoyant power in most cases. Suppose further that no obvious interfering factors, such as subjects’ being distracted, drunk, tired, etc., are observed in those experiments that can explain the subjects’ failures, and as a result ectoplasmic interferences are appealed to. The appeal to ectoplasmic factors, however, obviously has the following problem. We do not have a respected scientific theory about ectoplasmism in which other phenomena than the failures in clairvoyance experiments are explained. Ectoplasmic factors thus seem to be a mere concealed way of tautologizing the law about clairvoyance. The requirement in Pietroski & Rey’s proposal that interfering factors be “independent” thus means that those factors not only explain the

occurrence of counterexamples to the law under scrutiny, but they can also do some explanatory work in some independently vindicated theories. Hence, if counterexamples to a cp-law can be explained by independent interfering factors in the above sense, the cp-law is not vacuous. The clairvoyance case shows that if factors cited to explain counterexamples to a cp-law are subsumed under ill-warranted theories or under no theories, then the cp-law has to be regarded as vacuous. On the contrary, if counterexamples to a cp-law can be explained by factors described by well-warranted and substantive theories, then the cp-law is non-vacuous.

Pietroski & Rey, in my view, has offered a plausible sufficient condition for the non-vacuity of cp-laws. The main idea is that cp-clauses or provisos carry with them promises or commitments that counterexamples to a theory can be explained by (mostly) other theories that are themselves substantial and well warranted. If it turns out that no such theories are available after our repeated efforts to find them, then the theory with a proviso or cp-clause attached would be construed as genuinely vacuous. We would then disregard the theory.

Pietroski & Rey’s proposal has an important bearing on solving KW’s error problem for a realist dispositional account of rule-following, which, as we have seen, is fundamentally a problem of vacuity for a proviso or cp-clause. It can be solved as follows from Pietroski & Rey’s view. Insofar as interfering factors cited to explain a subject’s “deviant” rule-following behaviors are subsumed under a substantive and warranted theory, the hypothesis ascribing a real disposition to the subject should not be charged with vacuity or circularity. For example, a subject may answer “5” or “122” or another number when queried “68 + 57 = ?”. The fact that the subject has been hit on his head can be reasonably cited to explain why those performances are “deviant” or “erroneous” relative to a real disposition to follow the plus rule. Meanwhile, substantive and warranted theories of neurophysiology are available to explain how those deviant performances of calculation may occur when one’s head was hit, in addition to explaining many other phenomena. All other seeming counterexamples to a disposition-to-follow-a-rule hypothesis (i.e., seemingly “erroneous” behaviors) can be treated similarly. Namely, interfering factors will be cited to account for their apparent deviation from the underlying disposition and they must be

---

5  P & R also apply their proposal to KW’s error problem. See especially section 3.3 of P & R (1995). My aim here is to pursue and elaborate on their idea.
subsumed under well warranted theories in order for a disposition-ascribing hypothesis to be non-vacuous. On the other hand, if repeated efforts to find an interfering factor, which is subsumed under a substantive and warranted theory, have failed, the “deviant” behavior would have to be regarded as a manifestation of a disposition that differs from the one hypothesized. The original hypothesis then should be given up in favor of another.

Pietroski & Rey’s “cheque-cashing” proposal then makes perfect sense within Martin & Heil’s realist web picture of dispositions. A proviso or cp-clause is inserted in a disposition-ascribing hypothesis to deal with counterexamples (i.e., “erroneous” behaviors that deviate from the hypothesis’ predictions). Adding a cp-clause or a proviso, according to Pietroski & Rey’s proposal, amounts to making an explanatory commitment that there exist interfering factors causing those deviations, and that there are or will be scientific disciplines to study how those interfering factors affect a subject’s rule-following behaviors. Studying interfering factors in various scientific domains is equivalent to uncovering the web-structure of the subject’s rule-following disposition. It is clear that Martin & Heil offer a plausible metaphysical account of dispositions, and Pietroski & Rey provide a sound methodology with which the empirical reality of dispositions can be investigated in a scientifically respectable way. The result is that a subject’s real disposition, claimed to constitute his following a particular rule, can be determined non-vacuously and non-circularly. Consequently, the error problem that emerges in Martin & Heil’s realist dispositional account of rule-following is satisfactorily solved.

How might Pietroski & Rey’s account fare if KW’s skeptic insists that cp-clauses and provisos are hopelessly vacuous, or, “hedged”, that no sufficient condition for non-vacuity can be offered, including Pietroski & Rey’s proposal? Indeed, KW rejects any theory or hypothesis that employs a cp-clause or proviso to ascribe a disposition. However, the problem with rejection of a cp-clause or proviso is that KW would have to face the undesirable consequence that he has to call into doubt almost all respectable sciences, including physics. For not only do special sciences, such as psychology, linguistics, economics, and so on, rely on cp-clauses or provisos extensively to establish causal laws, but very basic physical theories, such as Newton’s theory of gravitation, theory of magnetism, Boyle’s theory of gases, etc., also implicitly embody provisos or cp-clauses. There is serious doubt whether there is any current fundamental science that is
free of cp-clauses or provisos. Even if there were, they would be extremely rare. As a result, if KW’s skeptic rejects the legitimacy of cp-laws in the practice of science, he will inevitably rule out almost all of our knowledge about the world. This consequence renders KW’s reply seriously deficient. Also, it trivializes KW’s position as a form of inductive skepticism.

KW would insist on the distinction between basic sciences and special sciences. One way he can do this is to make a distinction between a theory and its application. The theory of gravity *per se* has no cp-clause, for it only claims that any two masses attract each other in a certain way, period. It is only when the theory is used to predict the motion of specific bodies that a cp-clause becomes necessary. The distinction between a theory and its applications, KW would claim, does not seem to apply to theories in special sciences, especially in the branch of psychology that ascribes rule-following dispositions to people. If the distinction in question can be sustained, then the problem of vacuity of cp-laws only applies to special sciences but not to basic sciences. KW’s dismay at employing cp-clauses in psychological laws of rule-following would then revive.

We may well recognize that there exists the distinction in question between basic sciences and special sciences. Nonetheless, I would make two points in reply. First, even if we grant that only theories in special sciences employ cp-clauses, Pietroski & Rey have offered us a plausible way to formulate the sufficient condition for the non-vacuity of cp-theories. It would thus be methodologically legitimate to use those cp-theories that meet the non-vacuity criterion. Second, theories in special sciences, such as biology, linguistics, psychology, economics, and so on, do provide, at least sometimes, as good explanations and predictions for the human world as basic sciences for the physical world. It is not clear how KW might explain the explanatory and predictive power of theories in special sciences, unless he is able to recognize that theories in special sciences may indeed pick out nomologically significant properties in the world. Therefore, the laws or dispositions referred to in the psychology of rule-following may well be properties as real as those picked out by theories in basic sciences.

There is also a possible regress problem that KW’s skeptic would raise against

---

6 Cartwright is noted for her claim that “the laws of physics lie”. See N. Cartwright (1983), *How the Laws of Physics Lie*. 
Pietroski & Rey’s proposal. What if the independent theories under which the cited interfering factors are subsumed are themselves cp-theories? Will they depend on yet further theories, which might contain cp-clauses? Must we eventually get to theories that are wholly strict, not involving any cp-clauses? If there are indeed such theories, why don’t we just use them and jettison cp-theories?

Even if we grant KW’s worry that independent theories themselves contain cp-clauses and that they might in turn rely on further theories that also employ cp-clauses, my claim is that cp-theories remain methodologically sound. The reason is that cp-theories in a science network are mutually corroborative and coherent. A cp-theory cannot contradict the evidential results produced by other cp-theories. Otherwise, the credibility of the cp-theory would be undermined. Conversely, if a cp-theory receives and continues to receive evidential support from other cp-theories, then the cp-theory is very likely to be true. The same criterion applies to all other cp-theories as well. They themselves have to be coherent with other cp-theories in order to be supported and to gain credibility. This process of seeking corroboration in a knowledge network ensures that cp-theories are subject to critical and sophisticated examination. Consequently, the regress problem about cp-theories that arises in special sciences in general and rule-following in particular can be adequately handled.

But what if all theories contain cp-clauses? Would there be any independent confirmation? I tend to favor the view that at least some theories in basic sciences do not employ cp-clauses. As a result, cp-theories in special sciences can be ultimately grounded in them without running in an epistemic circle. As to the question of why we don’t simply jettison cp-theories and use theories in basic sciences, my answer is that there may be nomologically significant properties in the world that can be picked out only by cp-theories in special sciences, rather than by theories in basic sciences. As Fodor7 has forcefully argued, psychological properties are similar to, say, the monetary property in that they both cannot be described by vocabularies at the physical level. The monetary property can be realized by different physical material, such as papers, gold, diamond, etc. But then for any laws at the physical level to capture what all those possible realizations share in common, the laws will become too wildly

---

7 See Fodor, “Special Sciences, or The Disunity of Science as a Working Hypothesis,” in Block (1980).
disjunctive to be genuine laws. The idea to reduce all real properties to physical properties and to accordingly abandon cp-theories in special sciences in favor of theories in basic physical sciences may thus be simply a prejudice of positivist philosophy of science. To emphasize, I do not have to offer an explicit specification of a cp-clause in ascribing a rule-following disposition. All I have to do is to argue that a dispositional ascription and explanation can be non-vacuous. Since a suitable sufficient condition of non-vacuity can be offered, the epistemological difficulty in the error problem is resolved.

V. Solving the Finitude Problem

Underlying the finitude problem is the under-determination problem, which is a form of Humean skeptical problem about induction

To show how to solve the finitude problem, we have to start with Martin & Heil’s own solution to the problem. They argue that the finitude problem is an unfortunate result of KW’s misconception of the metaphysical nature of a disposition. A disposition, according to Martin & Heil, is a real property and infinity is built into it at the most basic level. For example, a quark of type A can come into causal contact with a potentially infinite number of type-B quarks for possible manifest behavior of a certain sort. A type-A quark can also interact with a type-B quark potentially at any time. Thus, a type-A quark has a disposition of an infinite nature in the sense that the disposition can project itself to anywhere and at any time for a possible manifestation. Similarly, a subject’s having a disposition P to follow a particular rule (e.g., to answer “4”, given the query “2 + 2 = ?”), or even, to use “+” to mean plus) is infinite in the sense that it can project itself to any time in the future in any problem case. A subject’s rule-following disposition may be dissimilar to a quark’s disposition to interact with a quark of another type in that the subject’s disposition can be repeatedly manifested while the latter disposition can perhaps be manifested only once. A quark’s disposition to interact with other quarks may be compared to the disposition of a bomb to explode: once the bomb is exploded, it loses its explosive power. On the other hand, a subject’s rule-following disposition is analogous to a bridge’s having a supporting dispositional power which is not lost after a car passes on it. Nevertheless, there is a clear sense of infinity in both cases: a subject’s or object’s disposition is projectible to an potential infinity of times for a manifestation under a potential infinity of input stimuli.
Although Martin & Heil’s view of dispositions is insightful and convincing, the Under-determination problem that underlies the finitude problem remains unresolved. For all the rule-following behaviors of a subject still under-determine the rule that he follows, even thought the subject has a real disposition that is ready to be manifested at any instant in the future. This problem is a sort of Humean inductive skepticism. A Humean skeptic about induction would ask us why, given the observation that the sun rose in the past, we have any reason or justification to believe that the sun will continue to rise tomorrow. In other words, why are we justified in projecting an observed regularity in the past onto the future and believing that the future will resemble the past? Such a skeptical doubt about induction underlies the skeptic’s doubt regarding whether it is recursion or recursion that is built into an agent’s disposition. If an agent has a disposition to recursively apply basic routines, then the agent’s future behaviors will resemble his past behaviors. In such a case, the agent would be following the plus rule. On the other hand, if an agent has a disposition to recursively apply basic routines, then the agent’s future behaviors will not resemble his past behaviors, and he would hence be following a non-standard rule. It is clear that underlying the finitude problem is a Humean problem about induction.

A reliabilist solution to the finitude problem

My overall response to the problem is to adopt a reliabilist position about the inductive skepticism involved. I do not plan to, and cannot, embark on the whole discussion of reliabilism. The main idea of reliabilism is that a subject’s belief that \( p \) is justified if \( p \) is produced by some reliable process of belief-formation. The subject does not have to be aware of the factors that are relevant to how he comes to possess the belief that \( p \) in order for his belief that \( p \) to be qualified as knowledge. In this regard, reliabilism is often called an “externalist” approach to knowledge.\(^8\) I believe that such an externalist view of knowledge (and justification) will offer us a proper response to KW’s skeptical view that every

use of a term or every act to follow a rule is merely “a leap in the dark.” Our intuitive idea, contrary to KW’s, is that we just know what we do when we follow a rule or mean something by a term. My claim is that reliabilism can do justice to our intuitive idea in question.

A reliabilist solution to the finitude problem is, roughly, this. When I watch you add, I form, on the basis of a finite sample of your behavior, a hypothesis about what infinitary disposition you possess inside you. However, KW’s skeptic would ask: How do I, and how can I, know this. My reliabilist response is: simple. We can, and do know what disposition a person has, and hence what rule the person is using, even if we ourselves cannot articulate a logically decisive reason for generalizing about the person’s disposition one way rather than any number of other ways logically consistent with the data. For so long as our tendency to generalize about the person’s disposition happens to be right, we are epistemically warranted in believing that the generalization is right. If such a reliabilist approach to the Humean skepticism about induction has a chance of succeeding, we would be justified in believing that a subject possesses a disposition, acquired through the interplay of his innate cognitive constraints and his past learning history, to apply basic routines recursively.

Let me elaborate the reliabilist solution a bit further. A standard form of reliabilism, forcefully advocated by A. Goldman (1979, 1986, 1988) and others, states that a belief is justified if and only if it is produced by a reliable cognitive process. A reliable process is one that is capable of producing a sufficiently high ratio of true beliefs. According to reliabilism, a subject need not know that a belief-forming process is reliable for the belief produced by that process to be justified. All that is required for a belief to be justified is that the belief-forming process is reliable. So, for example, you believe the sun will rise tomorrow, because of your tendency to project past regularities and the observation that the sun has risen in the past. You will be justified in believing that the sun will rise tomorrow, reliabilism says, if the sun will rise and if your belief-producing mechanism of inductive inference is reliable (i.e., if your inductive extrapolations will continue to be mostly correct). Similarly, since a competent adder will most likely apply basic routines repeatedly in the next new case and the belief-forming mechanism of inductive inference is reliable (if it has so far been very successful in producing true beliefs about a subject’s calculations and will continue to be so in the future), we are justified in accepting the inductive belief that people will repeatedly apply those basic routines they acquired through learning and therefore follow the recursive rule.
I do not intend to discuss reliabilism as a form of theory of knowledge and the theoretical difficulties it might face in any further detail. My aim is to point out that underlying KW’s finitude problem is a Humean, epistemological problem about extrapolation, and that a reliabilist theory of knowledge offers a way to provide the epistemological basis for the dispositional account of rule-following. The skeptical doubt about the finitude problem on a realist dispositional account of rule-following would consequently disappear.

References
---1977: Dispositional Explanation, in Dispositions, Tuomela, R. (ed.)