Article

Theory & Psychology

Retiring Popper: Critical realism, falsificationism, and the crisis of replication

Theory & Psychology
2024, Vol. 34(5) 561–584

DOI: 10.1177/09593543241250079
journals.sagepub.com/home/tap
Sage

Robert Archer
Independent Researcher

Abstract

The recent so-called crisis of replication continues to dominate psychology's methodological landscape. It is argued here that the apparent renaissance of Popperian thinking that characterises some of the key responses to the crisis of replication is fundamentally flawed. In essence, there is a serious lack of any sustained and rigorous treatment of ontology that underpins much of the current debate about replication and Popper's falsificationist approach. The overriding problem is that the replication debate reflects the methodologist tendency for mainstream psychologists to avoid or gloss over crucial ontological questions. In contradistinction, this article (a) underscores the primacy of ontology; (b) delineates and applies a critical realist stratified ontology to psychology; (c) utilises the latter as a springboard from which to argue for Popper's methodological "retirement"; and (d) revindicates the indispensability of context and the subtlety of psychological phenomena in arguing for the intrinsic limits of replication and experimentalism in general.

Keywords

context, critical realism, essentialism, falsificationism, Popper, replication

Maarten Derksen's (2019) recent article "Putting Popper to Work" provides the impetus behind my decision to write this article with the diametrically opposed title of "Retiring Popper." For Derksen, a "sweeping" feature of the reform movement that has developed in response to the so-called crisis of replication is the explicit use of the work of Karl Popper.[1] Continuing with his metaphor, my aim is to establish why Popper is the wrong philosophical "brush" and to provide a critical realist explanatory framework that accounts for the recurring failure to replicate in both psychology and the social sciences in general. Indubitably, the scientific rationale of replication is to ensure confidence in

Corresponding author:
Robert Archer, Independent Researcher, Parva, Street End Lane, Heathfield, TN21 8SA, UK.
Email: Robert.archer@warwickgrad.net

the development of new knowledge of causal mechanisms and processes. And it is precisely the trustworthiness of our findings that motivates further research and theoretical development. However, the failure to replicate has resulted in a so-called crisis that questions the very legitimacy of psychology as a scientific discipline. Indeed, as Derksen rightly notes, the irony of such replicative failure is to problematise the appropriateness of a methodology in which replication is deemed fundamental.

Basically, the problem centres on a failure both to address ontology and to examine its reciprocal relationship with methodology. Thus, whilst David Trafimow (2017), for example, maintains that for psychology to be a science, the method of reproducibility must be retained, critics have shown that the very reason for replicative failure is ontological, namely, the context-sensitivity, subtlety, and variability of human behaviour (Derksen & Morawski, 2022; Iso-Ahola, 2017; Van Geert & De Ruiter, 2022). However, the ontological critique of replication has not resulted in its methodological reevaluation. Instead, reformers have focused on refining methodology, with Popper "wheeled in" (or used as a "sweeping broom," to continue with Derksen's metaphor above) in defending the scientific status of psychology and the primacy of replication. In turn, this invites the charge of "methodologism," whereby matters of ontology are deemed of secondary importance or even irrelevant (see Teo, 2018).[2] In fact, Van Geert and De Ruiter (2022) have provided extensive evidence of methodologism, noting that the "dominant discourse regarding the necessary solutions to the reproducibility or replication crisis is one of *ever refined and stricter methodological recommendations* [emphasis added]" (p. 208).

Equally, the use of Popper vis-à-vis the legitimation and/or defence of replication is almost exclusively conducted at the levels of epistemology (what we can know, the nature of truth) and methodology (how research should proceed), without any substantive attention paid to ontology (what the world is like). There is no exploration and assessment of Popper's natural scientific ontology that underpins his normative requirement for direct replication and falsification, and no convincing defence of whether—Popper notwithstanding—direct replication is appropriate for psychology. The unfortunate consequence is that we can end up in seemingly inexplicable culs-de-sac. This arises because the nature of what we are researching has not been explicitly articulated and/or sufficiently grasped, which means our methodology has no firm grounding, and so we do not obtain the results we might expect from our research methods. It is important, then, that I provide a concise account of Popper's propositions germane to the replication crisis. After developing the critical realist framework, I will argue that Popper's natural-science ontology, ironically, renders impossible his insistence upon replication: put simply, Popper cannot do experimental science. This is due to his deeply flawed solution to the so-called *problem of induction*. My critique will help explain why he accords pride of place to deductive reasoning and the falsification (rather than confirmation) of scientific theories.

Furthermore, whilst this means Popper is not an appropriate ally for the reformers, I will argue that the continuing call for replication nevertheless remains inherently problematic for both the natural and social sciences because of the open-systemic nature of their respective subject matters. Moreover, in addition to the global open-systemic constraints on replication, I aim to sharpen, ontologically speaking, the psychological

critics' correct emphasis on sociocultural sensitivity and human variability, since it is the subtle nature of psychological phenomena that delimits experimentalism in general.

Critical realism and the strata of the real, the actual, and the empirical

The replication crisis has highlighted the longstanding issue of whether psychology can be modelled on the natural sciences, with the Popperian-inspired reformers voting for an unequivocal yes. Critical realism (Bhaskar, 1975/2008, 2015, 2016) argues for its qualified application in the social sciences.[3] Such qualification reflects the critics' arguments about the nature of psychosocial reality, which complement the nascent development of transcendental critical realism in psychology (see, e.g., Booker, 2021; Pilgrim, 2020; Willis, 2023) where Bhaskar's extension of his development of transcendental realism to the social and human sciences involves adopting different strategies and methods precisely because the human psychosocial world is fundamentally different from the physical world. Nevertheless, critical realism argues that psychological phenomena are such that they can be explained in fundamentally similar causal terms to natural phenomena. In assessing such similarity, we need to look at the early development of transcendental realism, which was concerned with the natural sciences.

In *A Realist Theory of Science*, Bhaskar (1975/2008) set out to reclaim ontology and establish a new non-Humean ontology. In undertaking the latter project, he employed a transcendental argument. That is, in asking what the world must be like for experimental activity to be possible, he showed that the world must be independently real and structured. For Bhaskar, the independent natural world necessarily has depth. Quintessential to his transcendental realist account of both the natural and social sciences is his metaphor of *ontological stratification*, which captures the fact that generative natural and social mechanisms[4] are irreducible to their actualisation.

The nature of experimental activity contra Hume

How does Bhaskar establish the transcendental indispensability of his stratified ontology? Well, for Bhaskar, the nature of experimental activity is completely at odds with the dominant Humean understanding of causality, which holds a scientific law to consist solely of observed "constant conjunctions" of atomistic (or unstructured) events. For Hume, the external world consists of nothing more than contingently related events, and so causality is untenably implied to reside in the (observed) invariant regularity of such events: instead of A caused B, we have A occurred followed by B over time. Thus, the basic problem with the Humean approach is its conflation of patterned regularity with natural necessity (i.e., with the causal powers and properties of natural objects). So, for example, every morning when I boil a kettle of water (event A) and make a fresh cup of coffee (event B), Hume cannot account for the fact that such regularity is explained by the irreducible natural properties of water, electricity, and so on, and the necessary connection between my purposeful activity of filling the kettle, switching it on, and so forth. Indeed, my watching the news on the television after I make my coffee every morning

(event C) shows just how untenable his position on causality is, since the observed constant conjunction of events A, B, and C precludes the fact that such events involve relations of both necessity and contingency (watching the television is not a necessary coffee-making precondition).

Basically, experimental activity is simply unintelligible in Humean terms. For in experimentally testing the boiling point of water, for instance, we are seeking to explain why water boils at 100 degrees Celsius, which is independent of, and irreducibly contingent to, any observed regularity. And it boils because the water molecules have enough kinetic energy to overcome their intermolecular bonds, break away from each other, and form gaseous molecules of vapour. For Bhaskar, it is such knowledge that provides us with our transcendental warrant for positing real generative causal mechanisms that underlie any observed events. As he puts it,

> That generative mechanisms must exist and sometimes act independently . . . and that they must be irreducible to the patterns of events they generate is presupposed by the intelligibility of experimental activity. But is up to it actual experiments to tell us what the mechanisms of nature are. (Bhaskar, 1975/2008, p. 42)

At the same time, it is the latter that underpins his stratification of reality into the strata of the real, the actual, and the empirical. All three strata are real but necessarily differentiated to capture the fact that structures and mechanisms (the real) are distinct from the events to which they give rise (the actual) and that such events are not necessarily experienced (the empirical). All three levels must not, therefore, be conflated or collapsed so that we can provide satisfactory explanatory accounts of scientific activity. Humean empiricism remains wedded to the level (or domain) of the actual, that is, at the level of observable events. Bhaskar has termed this *actualism*, which entails a denial of the real existence of underlying generative structures that account for things and/or events.

Furthermore, Bhaskar introduces the categories of *intransitive* and *transitive* to differentiate ontologically between the independent existence of the objects, mechanisms, and events of the natural world and the processes of conceptual scientific investigation. For Bhaskar, the errors of conflating the intransitive and transitive objects of science can be characterised as committing either the epistemic or ontic fallacies respectively. The *epistemic fallacy* is the view that statements about being can be reduced to or analysed in terms of statements about knowledge; the *ontic fallacy* is the view that knowledge is direct and unmediated.

Experimentation, causal tendencies, and the artificial creation of invariant regularity

A further problem with Humean actualism is the assumption of invariant regularity or *regularity determinism*. This is the idea that for every event y there is an event x or set of events $x_1 . . . x_n$ such that the latter are constantly conjoined. This presupposes that nature is a closed system, which is fundamentally not the case. The whole point of experimentation is to create (artificially) invariant regularities to understand what the natural world

is like outside the experimental setting because the natural world is an *open system*. To create artificial (systemic) closure, there are two conditions that must be in place. The first condition is an intrinsic one, which means there must be no change or qualitative variation in the object if causal mechanisms are to operate consistently. The second, extrinsic condition for closure concerns the relationship between the causal mechanisms and those of its external conditions that must remain constant to ensure regularity of outcome. In turn, when both conditions are met, this necessarily precludes the development of any novel sui generis properties.

So, if we return to the example above of the boiling point of water, the successful replication of any experiment(s) to establish its 100 degrees C boiling point derives from the fact that scientists will have ensured both conditions for closure have been met, namely, that there are no impurities in the molecular structure of water (intrinsic condition) and standard atmospheric pressure obtains (extrinsic condition). Bhaskar (1975/2008) himself gives the example of a simple electrical experiment designed to illustrate Ohm's Law. Here, the wiring of an electric circuit and the generation of an electric current would constitute what he calls *experimental production*; and the maintenance of appropriate resistance levels, and so on, would then constitute what he calls *experimental control*. The aim of an experiment is to activate and isolate a single mechanism and record its effects. Outside a closed system, these will normally be affected by the operations of other mechanisms, "either of the same or of different kinds, too, so that no unique relationship between the variables or precise description of the mode of operation of the mechanism will be possible" (Bhaskar, 1975/2008, p. 43).

In the open system that is nature, then, we find that water does not always boil at 100 degrees C, which explains (a) the transcendental intelligibility of experiments and (b) the use of ceteris paribus clauses or reference to law-likeness. Talk of "other things being equal" or law-likeness derives from the fact that we are not dealing with the existence of Humean constant conjunctions. In turn, this signals the importance of causal laws conceptualised as tendencies that dispose things to behave in certain ways, which may be exercised without being realised or manifest in any particular outcome. Indeed, the very reason for Bhaskar's tripartite distinction between the real, the actual, and the empirical derives from the fact that powers may (a) remain unexercised (power of water to dissolve salt); (b) exercised but unrealised (water does not boil at 100 degrees C because of impurities); or (c) realised but unperceived (water boiling as result of volcanic eruption with no-one to observe). That water does not always boil at 100 degrees C highlights its tendential dispositional nature. Moreover, it underscores the nonexistence of perfect regularities (Anjum & Mumford, 2018). Of course, the fact that water does generally boil at 100 degrees C means that we can talk of tendential (or imperfect) regularity, but not invariant regularity.

Emergence and essentialism

For Bhaskar, there are three senses of stratification. The first sense concerns the causal structures that give rise (tendentially) to events, which furnishes his distinction between the real and the actual. The second sense consists in the multitiered stratification of reality that is revealed in the creative historical development of science. Thus, the observable

properties of water are explained by a deeper level of reality that is described by the theory of atomic number and valency, which is further explained by the theory of electrons and atomic structure. The third sense is defined by the special form of multitiered stratification shown by *emergence*. Emergence concerns the ontologically irreducible powers or properties of objects. Critical realists often cite water as exemplary here, since its power to extinguish fire cannot be derived from its constituents (hydrogen and oxygen are highly inflammable). Equally, our capacity to feel, to think, and so forth, cannot be reduced to our brain chemistry: our emotions and thoughts are sui generis real. For realists, then, the emergent properties of water, emotions, and thinking exist at a different stratum from their constituents.

Furthermore, realism is necessarily committed to the doctrine of essentialism, which is the view that objects possess certain essential properties that make them one kind of thing rather than any other. At some basic level, we need the concept of essence to be able to discriminate between objects, for example, between water and wood or people and animals (Lowe, 2008). The rationale of science is to "dig deeper" as to what makes water precisely what it is, and here we know that its essence is H_2O. In understanding essences and their emergent powers and properties, realism emphasises the importance of distinguishing between internal and external relations. Thus, irreducible necessity characterises water, namely, the internal relationality of hydrogen and oxygen (they each necessarily presuppose each other, and water would not be what it is essentially without them). As we have seen with its tendential power to boil at 100 degrees C, externally or contingently related objects like impurities may interfere with the effects of the exercise of its powers, but they do not affect one another in their respective essential natures.

Extending transcendental realism to psychology

Transcendental realism and the social sciences

In extending transcendental realism to the social sciences, Bhaskar (2015, 2016) has argued that its applicability turns on the extent to which an independent analysis of the objects of social and psychological knowledge is consistent with his transcendental realist theory of science. Bhaskar argues that there are three ontological limits to such applicability: social structures, unlike natural structures, do not exist independently of (a) the activities they causally condition; (b) agents' conceptions of what they are doing in their daily activities; and (c) may be only relatively enduring. Furthermore, he underscores the fact that human agency may operate under unacknowledged conditions of action, create unintended consequences, require tacit skills, and involve unconscious motivations. Finally, he argues that social systems are radically open, thereby rendering impossible any attempts at experimental closure. (The actual extent and nature of such "radical" societal openness is a matter to which I will return later.)

As a condition of its applicability, a transcendental realist model of social science requires an intransitive object of study that is characterised by ontological depth, emergence, and contains causal mechanisms. The intransivity of social structures derives from the fact that they both preexist and presuppose human agency; ontological depth concerns their causal irreducibility (social structures possess emergent properties—or

essences—that are irreducible to people); and they are causal mechanisms precisely because they necessarily delimit agential activity. As with the essential nature of H_2O, we can talk of internal necessity vis-à-vis the essence of social structure. For example, a teacher presupposes a pupil, a headteacher presupposes teachers, and so on. Here, we are talking about internal relations between roles, ontologically distinct from the individual people who fill them and whom they causally affect. The teacher–pupil relation is an irreducible emergent property because the powers deposited within the role modify the powers of the individuals qua individuals. This modification arises from the combination of internally necessary social relations: a teacher cannot self-award an A-grade certificate, just as a pupil cannot revoke the decision of a national examination board. Such powers do not reside in the properties of individuals but in the social relations that presuppose such individuals for their enduring efficacy and mediation (Archer, 2002a).

Equally, the causal powers of social structure may be conceptualised in terms of causal tendencies (i.e., as conditioning human agency), evidenced by the fact that not all students regularly attend school. In other words, in the case of pupil absenteeism, whilst socioeconomic structures provide objective reasons for attending school (the need to obtain qualifications for future employment, potential status opportunities, and so on), such structures do not operate within a closed system nor determine in hydraulic fashion. Indeed, it is precisely the combination of the conditioning nature of socioeconomic systems and people's irreducible psycho-biographies that account for imperfect regularities. In strictly causal terms, there is no difference between natural and social reality, which is demonstrated by the reality of contingencies that interfere with the exercise of causal powers (water does not always boil at 100 degrees C and pupils do not always attend school or follow teachers' instructions). The ontological difference consists in the subjective mediation of objective human-made social structures: hydrogen and oxygen cannot subjectively weigh up the benefits of boiling at 100 degrees C in the way pupils can weigh up the structured benefits (or penalties) of (non)attendance. Clearly, then, it is important to emphasise that whilst Bhaskar rightly accentuates the ontological gap between irreducible social structures and individual persons—such that he contends social science ought to be an account of the former and psychological science an account of the latter—we must not lose sight of the fact that there must be considerable overlap (Groff, 2004), which means that psychologists cannot bracket-off (or give a post hoc nod towards) complex sui generis socioeconomic causal mechanisms that are operative in the open system that is society, just as social scientists cannot bracket-off irreducible psychological causal mechanisms.

Popper's methodological individualist repudiation of emergent properties

As we have seen in the above example of the teacher–pupil emergent property, social reality, like natural reality, is transcendentally characterised by ontological depth or stratification, since we are dealing with the causal irreducibility of social relations to people. The sui generis nature of emergent social properties (from the micro, to the meso and macro levels) necessarily transcends individual predicates, and so any insistence upon purely individualistic explanations for behaviour would remove the very rationale of sociology, thereby also undermining any adequate theorising of the psychology of

agential motivation. It is at this juncture that we can begin to understand just how deeply problematic Popper's ontology is, and hence why he is both an inappropriate ally for the reformers and for social scientific methodology in general. For, as I will elaborate in more detail later, his repudiation of essentialism in favour of nominalism means that he is unable to sustain an ontology of necessity, that is, of irreducible social and natural causal mechanisms. The task of critical realist social theory is to identify and theorise the essential properties of sociocultural systems and their agential mediation (Archer, 2002a), whereas for Popper,

> the task of social theory is to construct and to analyse our sociological models carefully in descriptive or nominalist terms, that is to say, *in terms of individuals*, of their attitudes, expectations, relations, etc.—a postulate which may be called "methodological individualism." (Popper, 1961/2004, p. 126)

Such "methodological individualism" nullifies ontological emergence at a stroke, by reducing the essential nature of emergent properties (such as the internal social relation between teacher and pupil) to individual properties. In nominalist fashion, Popper (1961/2004) asserts that when we engage in theoretical abstraction, we are liable to see "a kind of permanent ghost or essence" (p. 126) behind observable events rather than accepting that we are only dealing with concrete entities such as crowds or groups. So, for Popper, we should be talking—reductively—in terms of "people in uniform" in lieu of such abstract concepts as "army." Whilst permanency is not a defining characteristic of an essence, the salient point is that it is hard to see how Popper can avoid the charge of psychologism (Chalmers, 1985) in view of his repudiation of emergence. Ultimately, this means we can never explain why individuals behave in qualitatively different ways in different social contexts since the (transcendental) grounds for the latter have been ontologically removed. For example, when at school, a teacher is not exercising powers as husband or wife. In explaining the exercise of marital powers, necessarily one differentiates a different spatio-temporal setting (in the marital home, and so on). The fact that children behave in qualitatively different ways at different times (why Christopher undertakes his maths examination at 1 o'clock, watches television at home at 5 o'clock and follows the instructions of his scout leader at 7 o'clock) warrants the autonomy of social structure (Archer, 2002a). Without such autonomy, how do we account for the fact that Christopher is unable to take his maths exam during his scout meeting? Such behaviour cannot be reduced to statements about individuals per se. It is not the characteristics of Christopher as an individual that account for his behaviour in the three irreducibly emergent structural settings. Thus, Chalmers (1985) rightly takes Popper to task for his assertion that the valid aspect of psychologism lies in the fact that "it rightly insists that the 'behaviour' and the 'actions' of collectives, such as states or social groups, must be reduced to the behaviour and to the actions of human individuals" (Popper, 1966, p. 87). Indeed, Christopher's scout meeting is not simply a social group composed of individuals who happen to hold specific beliefs about scouting, since such beliefs presuppose an irreducible emergent context as a necessary condition of their intelligibility.

Transcendental realism and psychology

For Martin et al. (2010), unless it can be demonstrated that at least some features of human psychology are irreducible to physical, biological, and sociocultural properties, psychology has "no distinctive subject matter of its own and can readily be absorbed by fields of inquiry judged more fundamental to the constitution of psychological subject matter (e.g., neurophysiology, evolutionary biology, computational science, and cultural studies)" (p. 9). Here, transcendental realism can provide the ontological underpinning that secures the distinctiveness of psychology. As we have seen vis-à-vis social science, a transcendental realist model of psychology requires an intransitive object of study that is characterised by ontological depth, emergence, and contains causal mechanisms. Here, Bhaskar (2015) maintains that intentional human actions are the result of the mental states of actors, and such mental states are the intransitive reasons that (transcendentally) constitute the psychological analogue of generative mechanisms in nature. Ontological depth concerns the irreducibly emergent causal efficacy of reasons, and they are causal mechanisms precisely because they explain human intentionality.

However, Bhaskar commits two mistakes in how he theorises intentional processes: (1) he views reasons as efficient causes and (2) he conceives of action as a subtype of physical event rather than as the execution of purpose (Groff, 2004). For Groff, reasons are not efficient causes but final (i.e., purposeful) causes. She points out that she is not describing the same mistake in two different ways, since (1) and (2) concern Bhaskar's incorrect answers to the questions: "In what sense are reasons causally efficacious?" and "What are actions?," respectively. In terms of (1), Groff is making the simple yet crucial point that our reasons for action are irreducibly meaningful and cannot be conceptualised in terms of efficient causation. However, in terms of (2), she is making a more complicated point that Bhaskar does not distinguish adequately between action and behaviour. Briefly, Groff draws upon the work of Charles Taylor (1985), who distinguishes between "causal" and "qualitative" ways of understanding human action. The former considers actions as being events like any others, which happen to have psychological causes and which are physical with nonmaterial causes; the latter holds actions as differing ontologically from other events, where actions and their causes cannot be disaggregated—they "just *are* the expressions of purposes; they cannot be characterised first in non-purposive terms, and only then. . . connected to a mental or emotional state" (Groff, 2004, p. 126). The problem is that Bhaskar untenably proffers the "causal" view, arguing that actions can be redescribed independently of their reasons.

The transcendental realist case against Popper's falsificationism

At the outset, I want to make clear that I find Popper's "critical rationalism" frustratingly contradictory and inconsistent, which I aim to establish in my critical discussion of his writings about scientific laws and essentialism. For example, he readily employs such realist terms as causality, knowledge, growth, and facts, and upholds the metaphysical realist correspondence theory of truth (Popper, 1972, p. 46).[5] Equally, he concedes that scientific confirmation is possible rather than theories ever subject to interminable

refutation (Popper & Eccles, 1977/1985, p. 47). However, we have already seen that his social ontology is ontologically "flat" in his adoption of methodological individualism, in turn robbing us of our ability to explain individual and collective social behaviour. I will now show that, despite his so-called "modified essentialism," he consistently adopts a nonstratified natural ontology, which has the unfortunate yet ineluctable consequence of removing the very basis of experimental science.

The false problem of induction, Popper's "modified essentialism," and the inevitable failure of deductivism

Popper's (1969) critical rationalism begins by rejecting induction as a scientific method and replacing it with a continuous process of conjecture and refutation. It is important, therefore, to spell out why the "problem of induction" is a reductio ad absurdum because its ostensible insolubility provides the basis for Popper's falsificationist deductivism. Basically, the Humean problem of induction concerns our alleged inability to assume that because an event (or series of events) has always been observed to occur in the past, it will do so in the future (e.g., because the sun has always risen in the past, it will do so tomorrow). As we have already seen, Hume proscribes making any causal inference by purely a priori means. Instead, in actualist fashion, he maintains it can only be based on observed constant conjunction. But this unavoidably forces us to reach the absurd conclusion that we lack any firm grounds for trusting all past experiences (the sun may not rise), and thus throws into sharp relief the fact we cannot avoid the transcendental realist argument for natural necessity, that is, the transcendental presumption of irreducible intransitive causal mechanisms. Of course, whilst we know the Sun appears to rise from the horizon (and that it is the Earth's motion that causes the latter), the crucial point is that we know that the hydrogen fuelling the sun is not likely to run out for a few billion years. In other words, induction is a false problem because science is in the explanatory business of discovering the intransitive nature and structure of the world (and the solar system), with the latter providing us with our rational warrant for trusting that the sun will continue to rise.

Ultimately, Hume's problem of induction is easily resolved by accepting a realist (stratified) ontology of internally structured and differentiated objects that possess causal powers and liabilities. In turn, we can then distinguish between qualitative change and mere successions of events and thus between necessary or causal changes and relationships and accidental ones (Sayer, 2010). This stands in stark contrast to Hume's universalisation of contingency, which is the sine qua non of constant conjunction. However, Popper is contradictory in his approach to natural necessity, essentialism, and natural laws. He stated that universal laws "explain regularities or similarities of individual things or singular facts or events," since they are "conceived . . . as (conjecturally) descriptions of the structural properties of nature—of our world itself" (Popper, 1972, pp. 195–196). Furthermore, "the 'principle of the uniformity of nature' can again be regarded as a metaphysical interpretation of a methodological rule – like its near relative, the 'law of causality'" (Popper, 1959, p. 253). Finally, the principle "expresses the metaphysical faith in the existence of regularities in our world (a faith which I share, and

without which practical action is hardly conceivable)" (Popper, 1959, p. 252). I cannot imagine any scientist, let alone critical realist, baulking at such talk of explanation, structural properties, and the principle of the law of causality.

Nevertheless, one can readily "pin down" Popper's ultimate position, which irredeemably undermines his talk of explanation of regularities in terms of their (necessary) structural provenance, at the same time removing the very possibility of doing experimental science:

> I reject all *what-is questions*: questions asking what a thing is, what is its essence, or its true nature. For we must give up the view, characteristic of essentialism, that in every single thing there is an essence, an inherent principle . . . which necessarily causes is to be what it is, and thus to act as it does . . .. We must give up the view that . . . it is the essential properties inherent *in each individual or singular thing* which may be appealed to as the explanation of this thing's behaviour. (Popper, 1972, p. 195)

Furthermore, Popper (1959) abstains "from arguing for or against faith in the existence of regularities in our world" (p. 253), and, moreover, he asserts that we cannot "*ever* [emphasis added] find out if any given non-logical statement that it is in fact naturally necessary" (p. 454). In fact, for Popper, the concept *necessary* is a "mere word – a label useful for distinguishing *the universality of laws* [emphasis added] from 'accidental' universality . . . I largely agree with the spirit of Wittgenstein's paraphrase of Hume: 'A necessity for one thing to happen because another has happened does not exist. *There is only logical necessity* [emphasis added]'" (p. 460).

As I said above, Popper's work is frustrating to read, and his equivocal "largely agree" vis-à-vis the Humean repudiation of natural necessity exemplifies this, since the point, of course, is that one either does or does not endorse natural necessity. In this regard, Popper's unequivocal rejection of "what-is" questions (essentialism) must surely be taken at face value, together with his assertion that *necessary* is a mere word, which incontrovertibly means he completely agrees with Wittgenstein and Hume's denial of natural necessity. In fact, Popper's response to a reviewer who described his position as a "modified essentialism" is clear in its antiessentialism: "to avoid misunderstanding, I wish to say here that my acceptance of this term [modified essentialism] should not be construed as a concession to the doctrine of essentialist definitions. I fully adhere to the criticism of this doctrine . . . given in my *Open Society, Vol. ii*" (Popper, 1972, p. 195). And in his *Open Society*, Vol. 2, he untenably asserts that

> methodological nominalism is nowadays fairly generally accepted in the natural sciences. The problems of the social sciences, on the other hand, are still for the most part treated by essentialist methods. This is, in my opinion, one of the main reasons for their backwardness. (Popper, 1966, p. 30)

This is a clear and unequivocal rejection of natural necessity, since science is not in the nominalist business of explaining nature: the essential properties of water should be sufficient to underscore this.

Crucially, then, his denial of natural necessity and essentialism in favour of nominalism necessarily undermines his insistence that we accept the methodological rule "that we are not to abandon the search for universal laws . . . nor ever give up our attempts to explain causally any kind of event we describe" (Popper, 1959, p. 61). For Boulter (2002), Popper thus wants to have "intelligibility on the cheap: he wants to join the empiricists in their anti-necessitarianism, and yet keep all the benefits just the same" (p. 73), which, of course, he cannot do. Indeed, it strikes me that Popper simply wants to have his (ontological) cake and (methodologically) eat it. And in eating his ontological cake, he ends up removing any rational warrant for scientific experimentation and discovery. The nonexistence of such a warrant derives from his repudiation of natural necessity and essentialist "what-is" questions (in favour of logical necessity) because an antiessentialist (i.e., nonstratified) ontology means we have nothing to get our scientific teeth into (given that irreducible natural entities are held no longer to exist). This is precisely what justifies Maxwell's (1972) argument that it remains completely problematic how we could ever have good grounds for supposing a theory to be falsified, and Sayer's (2010) point that Popper is forced to commit himself to a universalised ontology of contingency.

The inevitable failure of Popper's deductivist "solution"

Given that Popper will only countenance logical necessity, it is inevitable that his "solution" to the Humean account of the problem of induction is doomed to fail, since he has removed the ontological rug from underneath his scientific feet. The salient point is that logic does not, and cannot, exhaust reality. For the possibility of argumentation presupposes that logical relations among propositions and conclusions are only part of the story: we cannot make sense of any proposition of an argument if such a proposition is not ontologically grounded, that is, transcendentally it must be about something other than itself. This is precisely the problem: for anything to be falsifiable presupposes that there are correct irreducible ontological grounds for us to be able to falsify in the first place. However, central to Popper's solution is his "hypothetico–deductive" procedure in which scientists are held to advance "bold" hypotheses or conjectures from which testable propositions can be deduced. Here, Popper claims that by construing science as an attempt to falsify a hypothesis rather than verify it, we shift our focus from the generation of the hypothesis to the process of testing it (Burns, 2009). Essentially, Popper (1969) says we make a deductive prediction from our hypothesis and look for empirical evidence that falsifies the prediction. Yet herein lies the overarching problem: shifting our focus away from hypothesis-generation, that is, from addressing ontological "what-is" questions, results in a truncated or *logicist* philosophy of science that "has little penetration of what most theoretical disputes are about and is unable to say why we don't seek to fit just *any* data into such deductive systems" (Sayer, 2010, p. 114).

Of course, this is to be expected given that Popper denies natural necessity. But the upshot of his deductivist approach is that any conclusion can be deduced from an infinite number of (potentially absurd) premises precisely because there are no ontological grounds both to make sense of, and delimit, such premises. Hence Sayer's (2010) absurd deductive example of "All woods conduct electricity; copper is a wood, therefore copper

conducts electricity" (p. 115). So, whilst Haig (2022a) is quite right to note that any replicative failure would not threaten the inferential structure of the hypothetico–deductive model itself, this is to miss the point, for the model reductively redirects our attention away from natural necessity.

Time to dispense with falsificationism and reclaim the search for truth

I agree with Dykes (1999) that one of the most troubling aspects of Popper's philosophy is his refusal to countenance anything positive, despite his (inconsistent) acknowledgement that many theories are positively confirmed, that is, not falsified. Again, however, we should not be at all surprised that Popper champions falsification, since his rejection of natural necessity enjoins that there is nothing with which our truth-claims can correspond and explains his disdain for definitions (see Note 5). With consistency, then, he maintains thus:

> The way in which knowledge progresses, and especially our scientific knowledge, is by unjustified (and unjustifiable) anticipations, by guesses, by tentative solutions to our problems, by *conjectures*. These conjectures are controlled by criticism; that is, by attempted *refutations* . . . They may survive these tests; but they can never be positively justified: they can be established neither as certainly true nor even as "probable." (Popper, 1969, p. xi)

But knowledge cannot progress if all we are doing is conjecturing and refuting and our conjectures can never be positively justified. Of course, Popper is quite right to assert the latter, with his additional denial of the possibility of probability, for there is nothing to ground such nonprobable conjectures.[6] Whilst Popper (1972) commits a performative contradiction when he maintains that "all knowledge is hypothetical" (p. 30), since the latter proposition cannot be conjectural, ultimately, Dykes' (1999) correct point that Popper does not seem to understand that "truth implies a *'what is'* question every time critical rationalism tells *what is not*" (p. 14) stems from his illicit ontological destratification of the intransitive. It is the search for how things really are, which provides the rationale for the natural and social sciences. The Popperian negative injunction to falsify is thus wholly indefensible for transcendental realist reasons: first, the very practice of science is incomprehensible without the presupposition of an ontologically stratified natural world; second, truth is a precondition of all human inquiry, and is best understood as a formal regulative norm (Archer, 2002b).

The DREI(C) and RRREI(C) alternatives to falsification

In contradistinction to Popper's falsificationist hypothetico–deductive approach, Bhaskar (2016) proffers what he calls the DREI(C) schema, which involves a five-step logic of scientific discovery. Step one, D, consists in the *description* of some pattern of events or phenomena. Step two, R, involves *retroduction*, which involves the hypothesising and identification of the causal mechanisms that tendentially give rise to events and phenomena. Step three, E, requires the *elimination* of those mechanisms that do not apply. What follows next, step four, I, is the *identification* of the causally efficacious generative

mechanism or structure at work. Finally, step five, (C), involves the iterative *correction* of earlier findings in the light of this identification. For areas of research that are already well theorised, retrodiction, RRREI(C), may be utilised, which stands for *resolve* using existing theory, *redescribe* in terms of relevant theories, *retrodict* back to antecedent states of affairs; *identify* causal mechanisms, and possibly *correct* the overall picture considering the fuller explanation (Price & Martin, 2018).

Interrogating regularity and the intrinsic limits of replication[7]

My aim now is to show precisely why, as Stroebe and Strack (2014) argue, exact (or direct) replication is illusory. Fundamentally, pace Stroebe and Strack (2014), the alleged "crisis of replicability" is not primarily due to an *epistemological* misunderstanding but is attributable to an *ontological* one:[8] the issue is about the nature of causal mechanisms, their irreducibility, and relative interplay. My argument is that the very reasons for replicative and nonreplicative experimentation (paradoxically) disclose why we can never assume that we will always achieve consistency of results. In other words, the scientific norm of replication requires the artificial establishment of systemic closure—that is, Hume's "constant conjunctions"—but such artificiality necessarily enjoins degrees of "inexactness," since what happens in concrete real-world settings cannot be mirrored in controlled experiments. In other words, we cannot capture in isomorphic fashion natural necessity precisely because the latter is intrinsically open-systemic, which means some degree of inexactness is ineluctable. Indeed, as Anjum and Mumford (2018) note, even where scientists can create a limited and highly controlled context in a laboratory, a very small difference in experimental set-up might produce a vast difference in outcome. Again, this is due to the attempted artificial isolation of causal mechanisms in which real-world causal complexity is deconcretised, in turn making successful replication inherently problematic as crucial causal mechanisms may inevitably get missed.

Consequently, Anjum and Mumford (2018) argue that direct replication holds very little power and is a poor substitute for real reproducibility. They maintain that it is more important to see whether results can be reproduced in different ways than in the original study to assess the robustness and generalisability of a theory. In this regard, it is instructive that the Committee on Reproducibility and Replicability in Science (National Academies of Sciences, Engineering, and Medicine, 2019) found that:

> Non-replicability is a normal part of the scientific process and can be due to the intrinsic variation and complexity of nature, the scope of current scientific knowledge, and the limits of current technologies. Highly surprising and unexpected results are often not replicated by other researchers . . . *non-replicability of results is a normal consequence of studying complex systems with imperfect knowledge and tools* [emphasis added]. (pp. 85–86)

This underpins Anjum and Mumford's (2018) argument that natural reality is intrinsically "messy," since there are "many different processes under way, often interfering with each other's natural passage, and with unpredictable and indeterministic events occasionally mixed in" (p. 252). So, whilst the Committee is not calling for

the abandonment of replication, such "messiness" enjoins that scientists temper their expectations about replicative success. Here, the DREI(C) schema would take the normative "pressure" out of replication, since the (implicit) closed-system expectation of complete consistency is replaced with the open-system focus on the iterative correction of earlier findings.

Tendencies and "demi-regularities"

To recap, it is because natural generative mechanisms necessarily operate as tendencies (in concrete open systems, which causally depend on other mechanisms that cannot/do not get included in any experimental set-up) that accounts for the nonexistence of perfect regularities. And, of course, the existence of "imperfect regularities" characterises the social sciences (including psychology) as they, too, deal with open systems. However, there is a generic lack of clarity vis-à-vis the precise nature of regularity in the replication literature, some of which strikes me as perilously close to being parasitic upon a closed-system (i.e., nontendential) ontology. For example, whilst Haig (2022a) is right that, in the context of his discussion of Popper's insistence upon experimental closure, intersubjective testing practices do not have to take the form of replication, his (unqualified) comment that "the regularity of the occurrence of events underwrites replication and other reliabilist validation strategies" (p. 227) is redolent of presupposing systemic closure, since it is not the regularity of events, pace Haig, that underwrites replication, but rather the artificial creation of such regularity of events through human manipulation. Indeed, he states that "phenomena are *stable* [emphasis added], recurrent features of the world that we seek to explain. Many of them are *noteworthy regularities* [emphasis added] and are often called *effects*" (p. 225). The problem here is that empirical effects are not characterised by invariant "stability" or regularity: such effects are independent of Haig's "recurrent" stable phenomena. In other words, the psychological stability of agential decision-making processes (forethought, self-reflection, etc.), for example, is subject to open-systemic sociocultural contingencies that may interfere with or override such stable phenomena to produce irregular empirical effects (or none).

Of course, Haig might well retort that (real-world) regularity is not (and cannot be) invariant, and so we do need to be crystal clear about what we mean by regularity. First, then, the use of the stratified metaphor of the real, the actual, and empirical can help dispel any potential confusion, since Haig's "recurrent features" are not invariant at the levels of the actual and the empirical, and so we must be careful not to confuse or conflate the (relatively) invariant essential nature of mechanisms (level of the real) with their contingent manifestation in an open system (Haig's "effects"). Second, as Næss (2019) argues, we should differentiate nonclosed systems according to their *degree of openness*, rather than applying a strict dichotomy between closed and open systems,[9] in turn permitting talk of "demi-regularities" or "demi-regs" (which complements Anjum and Mumford's "imperfect regularities," 2018). Furthermore, he argues that demi-regs represent probability distributions existing in the real world. Such empirical demi-regs thus derive from relatively enduring causal tendencies (social structure). It is on this basis that Næss maintains that an understanding of demi-regs as repetitive in nature "implies that certain kinds of outcome (e.g. car commuting) are more probable in the given time-space

context than alternative outcomes (e.g. commuting by transit) for individuals belonging to a certain category (e.g. being a suburbanite)" (p. 477).

Næss's (2019) focus on probability relationships fits nicely with my discussion of the teacher–pupil emergent property, since the constellation of causal mechanisms here (education system, central government, economic relations, etc.) constitutes strong probability reasons for the empirical regularity of high pupil attendance. To reiterate, such probability relationships are not invariant—"they vary across time and space in occurrence as well as in strength. Still, in a given context . . . relatively strong relations may persist over a long time (yet with temporal variations in strength)" (p. 478). But we are still left with the task of explaining why some replications have been successful. Basically, such success derives from the fact that the structures of nature are differentially tendential because of their object-specific essential properties and simultaneous existence in an open system. As Anjum and Mumford (2020) argue, tendencies "come in degrees, just as a wineglass and a car windscreen are both fragile but the former to a higher degree than the latter . . . Some of those tendencies, especially in ideal conditions, will almost always manifest" (p. 117). I would suggest this can explain both why replicative success is possible (in some cases), and why such a large proportion of laboratory-based replications are reported to fail. That some generative mechanisms are more amenable to replicative success is attributable to their (object-specific) essential power(s) that underpin their durable stability and the extent of their known causal dependence on other mechanisms. In the case of glass, we are dealing with an invariant stable physical structure, which can be experimentally established by heating a mixture of lead oxide, zinc oxide, and boric acid strongly until it melts, and by changing, for example, the relative contribution of boric acid. We would reasonably expect a high replicative success rate in terms of how much boric acid is required.

Nevertheless, any experimental investigation of how glass actually "plays out" in the real world makes replication a trickier endeavour, since we find here that glass is not invariantly stable but is characterised by real-world degrees of stability (Gaviria et al., 2022). This is why experiments that test the strength of glass continue to remain tentatively exploratory (Santos et al., 2018) precisely because glass "inhabits" an open system where we cannot always know the extent of necessary dependence on other mechanisms and contingent interference. Thus, replication cannot be assumed to succeed on every occasion, as unique and unexpected causal factors may be discovered. However, this does not mean, of course, that we cannot say anything causally meaningful about the properties of glass. What we can say is that under certain conditions glass is more likely to behave in a certain way. It is in virtue of the nature of the object under consideration that we can talk of degrees of likelihood—hence why wine glass is more fragile than car windscreen glass as the former tends to shatter more easily than the latter at the concrete level.

It should be clear from the foregoing that natural science's endeavour to gain entrée to the open system of natural reality via the door of replication is inherently limited (which is why nonreplication has been found to be a normal part of the scientific process). It follows that replication is equally limited in the social sciences (including psychology) for the same open-systemic reasons. Despite this, it is not just the open-system ontology that underpins my serious reservations about Haig's (2022b, p. 822) assertion

that the social and behavioural sciences can close their systems of study sufficiently to allow for experimental practice, as well as with Stroebe's (2016) claim that exact replication is likely "the more a social psychological phenomenon is exclusively dependent on basic psychological processes" (p. 140). In addition, as I will argue, ontologically speaking, psychological phenomena do not possess the same kind of relative stability or invariance as physical phenomena because of their innate "plasticity." So, whilst I agree with Stroebe that there exist basic psychological processes (or "stable recurrent features," as per above), such processes are inherently susceptible to reductive misidentification if subject to any form of experimentalism (replicative or otherwise).

Revisiting the twin conditions for closure: The limits of experimentalism and the indispensability of (stratified) context

I now want to address my reservations about Haig's (2022b) claim that we can close psychosocial systems sufficiently to allow for experimental practice and that certain basic psychological processes can permit exact replication as Stroebe maintains. My argument centres on the fact that human beings are intrinsically open and irreducibly complex psycho-biological systems whose emergent psychological properties are ontologically distinct from the physical causal mechanisms studied by the natural sciences (Valsiner, 1987). Necessarily, and crucially, this means that the twin conditions for systemic closure can never always be met, thus delimiting exact replication and (artificial) experimental closure generically. As Sayer (2010) argues, we cannot adopt (closed system) experimentalism in the social sciences because "human actions characteristically modify the configuration of systems, thereby violating the extrinsic conditions for closure, while our capacity for learning and self-change violates the intrinsic condition" (p. 83). So, I agree with Haig (2022a) that it is plausible to suggest that psychology will reveal different rates of replicative success "with psychophysics tending to be high, social psychology considerably lower, and cognitive psychology somewhere in between" (p. 234). But whilst the latter clearly contradicts Haig's (2022b) claim that we can close psychosocial systems sufficiently to permit experimental practice, the question-begging issue remains precisely as to why we can expect such differential rates of success. In this regard, we need to address the twin conditions required for closure and develop Sayer's point originally made vis-à-vis the social sciences.

The intrinsic condition

I mentioned above that psychological phenomena do not possess the same kind of relative stability or invariance as physical phenomena because of their innate "plasticity" or malleability. The notion of plasticity concerns the intrinsic potential for personal systematic change over the course of a lifetime (Lerner, 1984). For such change to be possible requires (transcendentally) that psychological phenomena are not invariant or fixed, unlike the invariant properties of natural science, thereby undermining the intrinsic condition for closure. As Iso-Ahola (2017) neatly puts it,

> psychology is a science of subtleties in cognition, affect and behaviour. Its phenomena reside in and arise from the human mind, whether conscious *or non-conscious* [emphasis added], and as a result, are not static, but instead dynamic and changing, varying with internal and external conditions. (p. 2)

This means that any endeavour to effect experimental closure (replicative or otherwise) ever confronts unconscious mental phenomena, of which the relative causal efficacy cannot be predetermined. It is the latter that necessarily vitiates Stroebe's "likely" success and Haig's possibility of closure. It is hardly surprising then, as Haig himself admits, that any rate of replicative success in social psychology can be expected to be relatively low. Equally, the suggested relatively higher rate of success for cognitive psychology might be attributable more to de-concretised narrowness than to uncovering actual causality. But the salient point is that the "subtle nature of psychological phenomena and their sensitivity to social influences means that a previously observed effect may not appear under the seemingly same laboratory conditions at later times" (Iso-Ahola, 2017, p. 3). Such subtlety means they may not appear under the same conditions since there are unconscious influences at play, which cannot be intrinsically manipulated/predicted precisely because they are unconscious and inherently temporal and transient (De Luca Picione, 2021).

However, even though the intrinsic psychological condition for closure cannot generically be met (hence the relatively low level of replicative success in social psychology), this does not mean we cannot talk about causal tendencies and thus demi-regs in psychological terms. For whilst the gap between what we might call (following Anjum and Mumford) the "messy" reality of psychosocial life and the requirements of replication cannot be closed, it is not an unbridgeable one. Such variable "bridge-ability" derives from the degrees of stability that characterise psychosocial phenomena.

The extrinsic condition and the need for real-life (stratified) context

At the same time, the extrinsic condition for closure is equally problematic for replication and experimentalism in general. If we recall, the extrinsic condition for closure concerns the relationship between the causal mechanisms and those of its external conditions that must remain constant to ensure regularity of outcome. At the outset, the simple fact that any psychological experiment involves the complex interplay of conscious and unconscious psychic mechanisms means that the extrinsic condition can never be held a priori to be met, for there is an ever-present susceptibility to contingent psychic generative mechanisms. However, the main problem here concerns the fact that psychological phenomena are "more likely to surface in the real world than in artificial laboratory conditions" (Iso-Ahola, 2017, p. 3). This sums up the case against exact or direct replication and experimentalism in general, since psychological causal mechanisms qua tendencies are intrinsically intractable to abstracted isolation (as measurable variables) by virtue of their interdependent irreducibility, of which any adequate explanation enjoins spatiotemporal and historical contextualisation. This is not dissimilar to my example of glass experimentation, where data are established tentatively precisely because such experimentation necessarily disconnects or cuts-off key causal properties from their natural concrete setting.

In other words, as argued above, natural scientific experimentation unavoidably misses (to varying degrees) causal mechanisms and processes in effecting the extrinsic condition for closure. This applies equally to any psychological experiment since the latter requires an artificially created disconnection from concrete reality. The key point is that, as Van Geert and De Ruiter (2022) put it, "phenomena that researchers claim to control are therefore, in reality, actually entangled processes interacting within persons, in the context of materials, context and time . . . they cannot be manipulated independently of one another" (p. 210). This lack of contextualisation is consistently ignored, paid lip-service to, or even denied within the replication literature.[10] In essence, replication temporally freezes, ontologically flattens, and reductively decontextualises its subject matter.

To understand what we mean by interdependent irreducibility (i.e., context), critical realism begins with the need to understand the nature of abstraction and the importance of that from which we abstract (Sayer, 2010). An abstraction (or abstract concept) pinpoints a one-sided or partial aspect of an object. That from which we abstract is the multiplicity of features that constitute concrete objects such as people, social structures, and activities. As well as furnishing precision, the things to which abstractions refer have ontological parity with those things referred to by more concrete concepts. The concept of "concrete objects" denotes the fact that objects are constituted by a combination of diverse elements. As Sayer (2010) notes, a particular person qua concrete entity combines influences and properties from a wide range of sources (physique, intelligence, attitudes, and so on) "which might be isolated in thought by means of abstraction, as a first step towards conceptualizing their combined effect. . . . Abstractions should distinguish incidental from essential characteristics" (pp. 59–60). Furthermore, and crucially, sometimes the aspect of an object that is abstracted cannot exist in an abstract form "but *only in particular concrete forms* [emphasis added], which this 'contentless abstraction' ignores. . . [such contentless abstractions] seem harmless enough *until we come to put some 'explanatory weight' upon them* or *try to measure* [emphasis added] what is abstracted" (p. 67).

Intelligence, for instance, is a "contentless abstraction," which is why experiments are inappropriate to understand its nature because it can only be understood in concrete form, that is, in the context of the "real world." And it is this transcendental requirement that makes any endeavours to measure it doomed to failure because of its inherent concrete internal relationality. This is not to deny the sui generis reality of intelligence; it's simply to recognise that it is embedded within the complex irreducible psycho-biological systems of concrete embodied human beings who themselves are embedded in complex and open sociocultural systems. In experimentally "stripping" intelligence of the latter would be like trying to abstract and measure the nature of sugar in a cake. Or it would be like trying to measure the nature of "pupil effects" and "teacher effects" on a pupil's examination performance: the irreducible internal social relation between pupil and teacher (and its wider overdetermined psychosocial–systemic anchorage) cannot be measured in any meaningful sense. But this does not mean its effects, which are grounded in underlying social generative causal mechanisms, are not real. The latter point applies with equal force to understanding emotions, since they are internally related to cognition (Green, 1992) and so cannot be measured or isolated in experimentalist abstraction.

Furthermore, we have seen that cognitive psychology can be expected to be relatively more successful at replication vis-à-vis social psychology, yet cognitive scientists standardly employ abstract, as opposed to, concrete universals, which necessarily provides only half of the explanatory picture (Shillcock, 2014).

To be clear, then, the basic problem with both replication and experimentalism in general is the de-concretisation of psychosocial reality, thereby resulting in a fundamental loss of explanatory power. There is a lopsided focus on abstraction at the expense of the concrete and disregard of whether objects and processes are internally or contingently related: experimentalism "decontextualises those processes and controls *out* the complexity associated with real life" (Pilgrim, 2020, p. 26). It is the concrete (real-life) complexity to which Pilgrim refers that defines context. At the same time, such complexity encompasses the stratified and open nature of psychosocial reality, with no a priori limit placed on the number of nested social-cultural and psycho-embodied systems.

Concluding remarks

This article started with Derksen (2019) and now ends with him by responding to his conclusion. First, as he notes, the latest crisis in psychology has spawned a reform movement that is "distinctly Popperian" in nature, with the consequence that the reformers remain within "the bounds of a rather traditional conception of science" (p. 460). Quintessentially, such a traditional conception of science accords pride of place to experimental replication, held to be practised by natural scientists. However, it has been argued that Popper is a wholly inappropriate ally for the reformers primarily because of his antiessentialist ontology, which has the ineluctable consequence of removing the reproducible rug from under his methodological feet. Ultimately, Popper notwithstanding, the reformers would do well to take on board the fact that nonreplication is a normal part of the natural scientific process, which stems from the intrinsic difficulty of endeavouring to create systemic closure. As Anjum and Mumford have put it, natural scientists necessarily deal with a "messy" reality, and so we should not be surprised to discover such "messiness" in our open psychosocial world.

Indeed, and secondly, Derksen (2019) rightly underscores the ontological limit of the context-sensitivity of human beings and their changeable environment (the twin conditions required for closure) that makes replication in psychology an even more problematic endeavour compared with its natural-science counterpart (since the human psyche is not characterised by invariant stability). Third, and finally, Derksen thus queries whether some psychologists may decide to search for different approaches altogether, that is,

> away from quantitative methods and a search for causal laws. They will ask a question that is largely absent from the current crisis debate: What is psychology good for? And are quantitative methods and experiments always the best way to bring it about?" (p. 461).

Here, I have argued for a critical realist alternative, which entails a qualified application of Bhaskar's (1975/2008, 2015) transcendental realism. In contradistinction to Popper's falsificationist approach, a critical realist psychology rejects both replication and experimentalism and instead advocates the DREI(C) and RRREI(C) schemas, where

the exploratory identification of causal tendencies and mechanisms replaces the search for (invariant) causal laws.

Acknowledgements

I am grateful to the editor and especially Maarten Derksen for their very helpful comments and suggestions.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Robert Archer https://orcid.org/0000-0002-8253-4705

Notes

1. Derksen (2019) acknowledges that not "all aspects of Popper's work are represented in the discourse of the critics; nor that they always name Popper as the source of their ideas about replication, falsification, and criticism; nor that there aren't non-Popperian or even anti-Popperian elements in their ideas" (p. 457). However, whilst the explicit adoption of Popper's philosophy of science may sometimes be nebulous, unacknowledged, or denied, Derksen shows how the reformers are endeavouring to create a scientific practice that is congruent with Popper's philosophy of science in several key aspects.
2. Whilst Derksen (2019) does not use the term "methodologism," he is aware of the reformers' emphasis on methodological rules, and that the issue is simultaneously methodological and ontological.
3. The term *critical realism* combines Bhaskar's transcendental realism and his subsequent extension of transcendental realism to the social sciences, which he called *critical naturalism*.
4. I agree with Patomäki (2010) that whilst Bhaskar does not specify the exact meaning of "mechanism," arguably, it is not associated with mechanical forces, material-efficient causation, and the standard analogy of machinery. Instead, mechanism refers to "a thing that is capable of doing, or being acted upon, if it is triggered and not prevented by something else. Mechanism is thus both a wider category than force and a deeper category than law" (p. 67).
5. As Dykes (1999) argues, Popper's adoption of the correspondence theory is inconsistent with his disdain for definitions: "When we assert that a statement corresponds to the facts we mean that the words used accurately describe a specific external situation. But we could not assert correspondence if the words did not have precise meaning; i.e. did not have precise definitions" (p. 16).
6. Suárez (2013) underscores the noncausal nature of Popper's early "propensity theory of probability." However, Lawson (2008) argues that Popper's (1990) subsequent work on propensities (in which he maintains they are physically real and can produce an effect) should have led him to abandon deductivism because Popper (now) adopts the reality of forces and countervailing forces, which can shift. But, as Lawson himself points out, such propensities are not aspects of structures or objects, but of situations. Indeed, Lawson concedes that he is compelled to go "somewhat further" than the "later" Popper in his stratified portrayal of social reality.

7. For clarity, I am using the Committee on Reproducibility and Replicability in Science's (CRRS) definition of replication, namely, that to maintain that a study is replicable means that someone else can repeat it and obtain the same (consistent) results in answering the same scientific question (National Academies of Sciences, Engineering, and Medicine, 2019, p. 46). However, as Earp and Trafimow (2015, p. 5) note, whilst it has been possible to discern from a systematic review of the multidisciplinary literature 18 different types of replication, it is the distinction between *direct* and *conceptual* replication that shapes discussion in the field of psychology. Put simply, the former seeks to validate a particular fact or finding; the latter seeks to validate the underlying theory (cf. Haig, 2022a). But it strikes me that "conceptual replication" is a rather unhelpful and misleading misnomer, since its advocates rightly underscore the inherent stringent limitations (if not impossibility) of replication per se, thereby making it logically redundant.
8. Derksen and Morawski (2022) quote Stroebe and Strack's (2014) reference to "underlying mechanisms." The latter, however, is prefaced by Stroebe and Strack in terms of an "*epistemological* [emphasis added] misunderstanding" (p. 59). Of course, I completely endorse their talk of "underlying mechanisms," but a stratified, open-system ontology is not discernible in their article. In fact, they maintain that (a) direct replication is "less of a problem in studies where both the independent and the dependent variables are not culturally or socially mediated" (p. 61) and, moreover, (b) "the bad news is that theories cannot be verified or even demonstrated as probable" (p. 64). In terms of (a), as I will argue, replication is stringently limited in the natural sciences, and such stringency is even more problematic in psychology given that human psychology is inherently experimentally noncontrollable because of its context dependence and subtle nature; and in terms of (b), I have already argued that Popper's falsificationism is transcendentally false because of its antiessentialist ontology.
9. As already mentioned, Bhaskar (2015) has maintained that experimentalism is impossible because society is *radically* open. The danger of endorsing *radical* openness is that the social sciences are potentially stymied at the outset in terms of their ability to theorise relatively enduring causal tendencies. But empirically speaking, of course, this just is not the case and so we cannot assume a priori radical contingency.
10. As Van Bavel et al. (2016) note, "the role of context is frequently overlooked – and *even dismissed* [emphasis added] – in the evaluation of replication results. Several scientists have argued that hidden moderators such as context are *unlikely to influence* [emphasis added] the results of direct replications . . . others have argued that direct replications are the strongest (and possibly only) believable evidence for the reliability of an effect" (p. 6455).

References

Anjum, R. L., & Mumford, S. (2018). *Causation in science and the methods of scientific discovery*. Oxford University Press.

Anjum, R. L., & Mumford, S. (2020). Powers, probability and statistics. In A. S. Meincke (Ed.), *Dispositionalism* (pp. 115–130). Springer Nature.

Archer, R. (2002a). *Education policy and realist social theory*. Routledge.

Archer, R. (2002b). Reclaiming metaphysical truth for educational research. *British Journal of Educational Studies*, *50*(3), 339–362. https://doi.org/10.1111/1467-8527.t01-1-00207

Bhaskar, R. (2008). *A realist theory of science*. Routledge. (Original work published 1975)

Bhaskar, R. (2015). *The possibility of naturalism* (4th ed.). Routledge.

Bhaskar, R. (2016). *Enlightened common sense: The philosophy of critical realism*. Routledge.

Booker, R. (2021). A psychological perspective of agency and structure within critical realist theory: A specific application to the construct of self-efficacy. *Journal of Critical Realism*, *20*(3), 239–256. https://doi.org/10.1080/14767430.2021.1958281

Boulter, S. (2002). Hume on induction: A genuine problem or theology's Trojan Horse? *Philosophy*, *77*(1), 67–86. https://doi.org/10.1017/S0031819102000062

Burns, S. (2009). The problem of deduction: Hume's problem expanded. *Dialogue*, *52*(1), 26–30. https://www.academia.edu/809979/The_Problem_of_Deduction_Humes_Problem_Expanded

Chalmers, A. F. (1985). Methodological individualism: An incongruity in Popper's philosophy. In G. Curry & A. Musgrave (Eds.), *Popper and the human sciences* (pp. 73–87). Kluwer.

De Luca Picione, R. (2021). The dynamic nature of the human psyche and its three tenets: Normativity, liminality and resistance—Semiotic advances in direction of modal articulation sensemaking. *Human Arenas*, *4*, 279–293. https://doi.org/10.1007/s42087-021-00212-3

Derksen, M. (2019). Putting Popper to work. *Theory & Psychology*, *29*(4), 449–465. https://doi.org/10.1177/0959354319838343

Derksen, M., & Morawski, J. (2022). Kinds of replication: Examining the meanings of "conceptual replication" and "direct replication". *Perspectives on Psychological Science*, *17*(5), 1490–1505. https://doi.org/10.1177/17456916211041116

Dykes, N. (1999). Debunking Popper: A critique of Karl Popper's critical rationalism. *Reason Papers*, (24), 5–25. https://reasonpapers.com/pdf/24/rp_24_1.pdf

Earp, B., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology*, *6*, Article 621. https://doi.org/10.3389/fpsyg.2015.00621

Gaviria, J., Quijano, S., Quijano, J., & Ruiz, P. (2022). Computational study of the relative stability of some glass-ionomer cement-forming molecules. *Journal of Molecular Modelling*, *28*(10), 333–340. https://doi.org/10.1007/s00894-022-05211-x

Green, O. (1992). *The emotions: A philosophical theory*. Kluwer Academic.

Groff, R. (2004). *Critical realism, post-positivism and the possibility of knowledge*. Routledge.

Haig, B. (2022a). Understanding replication in a way that is true to science. *Review of General Psychology*, *26*(2), 224–240. https://doi.org/10.1177/10892680211046514

Haig, B. (2022b). Varieties of realism and their bearing on psychology. *Theory & Psychology*, *32*(5), 821–823. https://doi.org/10.1177/09593543211063840

Iso-Ahola, S. (2017). Reproducibility in psychological science: When do psychological phenomena exist? *Frontiers in Psychology*, *8*, Article 879. https://doi.org/10.3389/fpsyg.2017.00879

Lawson, T. (2008). Popper and social explanation. In T. Boylan & P. O'Gorman (Eds.), *Popper and economic methodology* (pp. 87–112). Routledge.

Lerner, R. (1984). *On the nature of human plasticity*. Cambridge University Press.

Lowe, E. J. (2008). Essentialism, metaphysical realism, and the errors of conceptualism. *Philosophia Scientiae*, *12*(1), 9–33. https://doi.org/10.4000/philosophiascientiae.222

Martin, J., Sugarman, J., & Hickinbottom, S. (2010). *Persons: Understanding psychological selfhood and agency*. Springer.

Maxwell, N. (1972). A critique of Popper's views on scientific method. *Philosophy of Science*, *39*(2), 131–152. https://doi.org/10.1086/288429

Næss, P. (2019). "Demi-regs," probabilism and partly closed systems. *Journal of Critical Realism*, *18*(5), 475–486. https://doi.org/10.1080/14767430.2019.1644951

National Academies of Sciences, Engineering, and Medicine. (2019). *Reproducibility and replicability in science*. The National Academies Press. https://doi.org/10.17226/25303

Patomäki, H. (2010). After critical realism? The relevance of contemporary science. *Journal of Critical Realism*, *9*(1), 59–88. https://doi.org/10.1558/jcr.v9i1.59

Pilgrim, D. (2020). *Critical realism for psychologists*. Routledge.

Popper, K. (1959). *The logic of scientific discovery*. Hutchison.

Popper, K. (1966). *The open society and its enemies* (Vol. II). Routledge.
Popper, K. (1969). *Conjectures and refutations: The growth of scientific knowledge*. Routledge.
Popper, K. (1972). *Objective knowledge: An evolutionary approach*. Clarendon Press.
Popper, K. (1990). *A world of propensities*. Thoemmes.
Popper, K. (2004). *The poverty of historicism*. Routledge. (Original work published 1961)
Popper, K., & Eccles, J. (1985). *The self and its brain* (2nd printing). Springer. (Original work published 1977)
Price, L., & Martin, L. (2018). Introduction to the special issue: Applied critical realism in the social sciences. *Journal of Critical Realism*, *17*(2), 89–96. https://doi.org/10.1080/14767430.2018.1468148
Santos, F. O., Louter, C., & Correia, J. (2018). Exploring thin glass strength test methodologies. In C. Louter, F. Bos, J. Belis, F. Veer, & R. Nijsse (Eds.), *Challenging glass 6: Conference on architectural and structural applications of glass* (pp. 713–724). TU Delft Open. https://doi.org/10.7480/cgc.6
Sayer, A. (2010). *Method in social science: A realist approach* (2nd rev. ed.). Routledge.
Shillcock, R. (2014). The concrete universal and cognitive science. *Axiomathes*, *24*(1), 63–80. https://doi.org/10.1007/s10516-013-9210-y
Stroebe, W. (2016). Are most published social psychological findings false? *Journal of Experimental Psychology*, *66*, 134–144. https://doi.org/10.1016/j.jesp.2015.09.017
Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, *9*(1), 59–71. https://doi.org/10.1177/1745691613514450
Suárez, M. (2013). Propensities and pragmatism. *The Journal of Philosophy*, *110*(2), 61–92. https://doi.org/10.5840/jphil2013110239
Taylor, C. (1985). *Human agency and language: Philosophical papers 1*. Cambridge University Press.
Teo, T. (2018). *Outline of theoretical psychology: Critical investigations*. Palgrave.
Trafimow, D. (2017). Commentary: Reproducibility in psychological science: When do psychological phenomena exist? *Frontiers in Psychology*, *8*, Article 918. https://doi.org/10.3389/fpsyg.2017.00918
Valsiner, J. (1987). *Culture and the development of the children's action: A theory of developmental psychology*. John Wiley & Sons.
Van Bavel, J., Mende-Siedlecki, P., Brady, W., & Reinero, D. (2016). Contexual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences*, *113*(23), 6454–6459. https://doi.org/10.1073/pnas.1521897113
Van Geert, P., & De Ruiter, N. (2022). *Toward a process approach in psychology*. Cambridge University Press.
Willis, M. (2023). Critical realism and qualitative research in psychology. *Qualitative Research in Psychology*, *20*(2), 265–288. https://doi.org/10.1080/14780887.2022.2157782

Author biography

Robert Archer started his academic career lecturing at the Universities of Birmingham and Bath, UK. His main publications are *Education Policy and Realist Social Theory* (Routledge) and, with Professor Martin Thrupp, *Education Management in Managerialist Times: Beyond the Textual Apologists* (McGraw Hill). His current research interests include psychosocial studies, philosophy of psychology, and social ontology.