

Prepublication Draft, for *Scientific Imperialism* , ed. Mäki, Walsh, Fernández Pinto

Logical Form, the First Person, and Naturalism About Psychology: The Case Against

Physicalist Imperialism

(Prepublication Draft, for *Scientific Imperialism* volume, ed. Uskali Mäki, Adrian Walsh, and Manuela Fernández Pinto)

Frederique Janssen-Lauret

University of Nottingham

1. Introduction: The Rise and Fall of the Physicalistic Empire

Physicalist philosophers and psychologists trust that all facts about human psychology will ultimately be explained in physical terms. The view that physics is such an exemplary science that it will eventually account for the psychological as well as the traditional domain of physics is a classic case of scientific imperialism as first defined by Dupré (1995). The explanatory capacity of physics, both with respect to its methods and to its domain, is taken to extend beyond the traditional realm of physics, and into that of psychology . But is it, as Dupré assumes imperialism must be, a harmful development? In the twentieth century, several practising psychologists modelled their methods and explanations on those of physics, preferring intersubjectively observable descriptions of behavior or brain states to first-personal self-reports. Assuming a naturalistic point of view, the physicalist approach to psychology and its philosophy seemed well-supported in view of these developments. At the time physicalism seemed like a case of neutral disciplinary expansion (see introduction, section 1.1) or even beneficial scientific imperialism (of the kind proposed in Mäki 2009). But, I will argue in this paper, more recent theories in psychology are moving in the opposite direction, and so naturalists ought to conclude that the project of physicalist imperialism has failed. As a philosopher of logic, I will focus on the under-appreciated problem of logical form: the recent move towards first-personal methods and theoretical statements in psychology, rather than impersonal ones modelled on the logical form of physics.

Twentieth-century philosophical naturalists had some good reasons to endorse physicalism, given the scientific practices of their contemporaries. At the time, psychologists strove to develop theories which shared a logical form with physics. Physicalist imperialism involved was therefore disciplinary, not domain-only (see introduction, 2.2). The logical form of physics is impersonal; it does not distinguish between first, second, or third person. All observations hold independently of who the observer is. Twentieth-century scientists equated this feature of physics to scientific objectivity, and strove to emulate it in psychology. They replaced first-personal self-ascriptions of psychological qualities with impersonal claims about behavior or brain states. From a naturalistic point of view, physicalism would have been vindicated if these efforts had led to improved scientific theories, theories which were simpler, more explanatory, more fecund, more predictive and testable. But a sample of the contemporary scientific literature in psychology reveals that practising scientists did not find the resulting theories fit for purpose. The twenty-first century saw a return to first-personal methods and theoretical statements, especially in areas like social psychology. As enthusiasm about physicalist imperialism has waned among empirical scientists in the relevant fields, twenty-first century philosophical naturalists should call for a re-examination of physicalism. I make the case that the recent resurgence of the first-personal in psychology points towards a new challenge for physicalism: the problem of expressive power. First-personal theories have a richer logical form, and potentially greater expressive power, than impersonal theories.

My case is based on philosophically interesting consequences of the difference in logical form between physics and psychology, and does not rely at all on familiar philosophical concerns about consciousness or the nature of the mind. A previously published argument of mine which uses only general principles belonging to the philosophy of logic, plus the plausible assumption that objects indiscriminable by physical descriptions are to be identified for the purposes of physics, implies that impersonal theories have diminished expressive power compared to theories with first-personal pronouns. Adding the equivalent of 'I' to a theory results in a minor but theoretically

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto
fruitful increase in expressive strength. Such theories can, where impersonal theories cannot,
discern objects which are descriptively indiscriminable.

2. Naturalism, Physicalism, and the Explanatory Scope of Physics

Scientific imperialism occurs when the explanatory scope of one discipline is stretched to encompass a domain traditionally assigned to another (called 'domain-only' imperialism in the introduction, 2.2) or when the methods of one science are forced onto another science which previously had its own distinct methodology. Physicalistic practitioners of psychology and its philosophy have tried, for philosophical reasons, to bring the methods and explanatory power of physics to bear upon psychology. I am concerned here with one previously-neglected way in which this physicalistic movement has been unsuccessful: a difference in logical form. Unlike most anti-physicalist arguments, the arguments of this paper do not flow from a priori philosophical opposition to physicalism. My point of departure is one many physicalists claim to share: philosophical naturalism. Before delving into the ramifications of the differences in logical form between impersonal physics and first-personal contemporary psychology, I will show in this section that there is conceptual space for a naturalist philosophy which is not physicalist. Subsequent sections will make the case that a non-physicalist naturalism is preferable to physicalist naturalism.

2.1 Naturalism does not imply physicalism

Naturalists need not be physicalists. Naturalism is a methodological thesis, which states that our best philosophical theories derive from fitting philosophy around the deliverances of the special sciences, and rejects first philosophy. Manuela Fernandez Pinto, in her paper in this volume, calls this type of naturalism the 'passive co-operation' type. As she puts it, naturalists of this type 'use the results of scientific research for limiting the scope of epistemological inquiry' (this volume, p. ??). I take a similar line, and have defended, in a joint paper, letting our best scientific theories guide our answers to questions of ontology and ideology too. We can arrive at credible answers to

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto

philosophical questions about the most general categories of things there are only by aiming to unify and clarify the theoretical statements of the special sciences into a coherent overall theory (MacBride and Janssen-Lauret 2015, p. 294). Naturalists admit no higher authority than our current best scientific theories, arrived at by means of what is presently our best scientific methodology. Physicalism, by contrast, is a metaphysical thesis, the philosophical doctrine that everything there is is physical. (One popular formulation of physicalism, 'everything is physical or supervenes upon the physical', is dubiously compatible with methodological naturalism. Supervenience requires quantification over possible worlds, purely philosophical entities rejected as posits by most current theories of physics, as we will see in section 6 below.) Naturalists should follow the scientific method where it leads, and endorse whatever our current best empirical theories are; they should be physicalists only if those theories talk only about physical things. Whether or not they do is a matter to be investigated, not stipulated.

Physicalism and naturalism are two logically independent doctrines. Physicalism does not entail naturalism; physicalists quite frequently base their physicalism explicitly upon reasons at odds with naturalism, such as philosophical intuitions or appeals to common sense. Appeals to common sense and intuition have at times been in conflict with, and been overthrown by, scientific progress. Cantor and Einstein did not let philosophical intuitions about the absurdity of infinite wholes the same size as some of their proper parts or non-Euclidean geometry stand in the way of the development of set theory and general relativity (MacBride and Janssen-Lauret 2015, p. 291). Nor does naturalism entail that everything is physical. Although it is perfectly possible that at some points in human history, our best evidence is in favor of physicalism, we must draw conclusions based on our *present* best theory. And, as I argue below, our present best theories of psychology are not physicalist. It cannot be assumed that the scientific method inexorably pushes towards the discovery that everything is physical. We cannot maintain physicalism based only on the hope that future science will confirm it. To assume physics will find a way to explain all facts about human

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto

psychology without actually having a credible way to explain them in such terms is not naturalism, but a kind of first philosophy, where intuitive philosophical convictions are allowed to take precedence over a posteriori empirical findings. We must opt instead for careful philosophical interpretation of recent scientific results, weaving our empirical findings into a unified theory without giving precedence to philosophical arguments from first principles. Demarcations of science from non-science ought to be made on methodological grounds. Maddy has argued that the very fecund and successful discipline of mathematics should not be denied the title of 'science' merely because its posits are for the most part not quantified over in physics (Maddy 1997). Similarly, naturalists should favor theories which score highly on our current best criteria of theory choice, whether physicalist or not. The correct naturalistic strategy is to investigate whether psychology can be interpreted in physicalist terms.

2.2 The Relationship Between Physics and Physicalism

Our best and fullest knowledge of physical objects derives from the science of physics, not from first philosophy. So naturalist philosophers who wish to be physicalists should derive the subject matter of their theories exclusively from the subject matter of physics, and model their methods as closely as possible on the methods of physics. Physics in itself remains neutral on the question of physicalism. Physics makes a variety of existence claims, quantifying over fields, particles, real numbers, space-time points, and other physical posits, but it does not state that its posits are all there is. Although it arguably quantifies over mathematical objects (Quine 1981, Putnam 1979), and assumes the truth of significant parts of mathematics, it does not say anything at all about the truth or falsehood, ontology or methodology of most other sciences. In particular, it says nothing about the sorts of objects discussed in the life sciences or social sciences, such as organisms, ecosystems, societies, or personalities, and therefore does not claim that they either fail to exist or are purely physical objects.

Where the evidence points towards all other sciences conforming to the methods and subject

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto

matter of physics, or to their discourse being reducible to that of physics, the methodological naturalist has good reason to embrace physicalism. The physicalistic project might have looked to naturalists like a harmless venture into disciplinary expansion at first, and there were indeed some good arguments for considering physicalist imperialism a viable project over the course of the twentieth century. The first is that practising scientists were making moves in this direction, especially towards explaining psychology in terms which referenced only impersonally observable behavior. Another derives from the search for a unified science (see also the introduction to this volume, section 3.3.) A unified science seems easier to arrive at if we only need to quantify over the domain of a well-established discipline like physics which we already accept as true, and consider the life sciences and social sciences as mere applications of it. As other disciplines, like biology and psychology, do contain much applied physics, physicalist disciplinary expansion was originally a promising idea. But where empirical results defy such treatment, it would be wrong to allow a philosophical programme, physicalism, to take precedence over careful weighing of a posteriori empirical findings. Unified science should result not from forcibly subsuming one discipline to another, but treating all theories which perform well on our best criteria of theory choice equally.

A unified scheme may just as well result from one scientific discipline adding to another. In this paper I offer one argument to suggest that this may be the relationship between current physics and current psychology. Twenty-first century psychology deploys first-personal methods and statements which suggest that psychology assumes the truth of physics, but adds further theoretical statements which are inexpressible in terms of physics' logical form. Several branches of contemporary psychology rely heavily upon statements in the first person, and theoretical statements about what social psychologists call 'the self', the referent of the first-person pronoun. Since the theoretical statements of physics are all paradigmatically impersonal, the methods of psychology differ from those of physics down to their very logical form. The canonical language of psychology does, where the language of physics does not, require grammatical differentiation

between self and other. The truly naturalist response is to search for the best interpretation of all deliverances of the empirical sciences, and consider the possibility that recent advances in first-personal psychology point towards a movement away from explanations in terms of physics. Of course it remains possible that physicalist will answer this challenge, and provide an interpretation of these results in physicalist terms. But the problem of logical form represents a new challenge for physicalism which must be addressed.

3. Impersonal Logical Form and Criteria of Identity in Physics

3.1 Impersonal Logical Form and the Regimented Language of Physics

The characteristic logical form of physics is entirely impersonal. Physics has no theoretical statements in the first or second person. Its methods, too, are impersonal. They are such that all evidence and observations must hold regardless of who the observer is. So the theoretical statements of physics correspond exactly to Quine's model of the canonical form of scientific theories: they have the logical form of classical first-order logic without individual constants. (Proponents of second-order logic may substitute 'second-order' for 'first-order' here. It makes no difference to my argument, which only pertains to the difference in expressive power between canonical languages with and without individual constants.) All physical pronouns are impersonal pronouns, the formal analogue of 'it': 'x', 'y', 'z', *etcetera*. Theoretical statements of physics have quantified logical forms like $\exists x\varphi(x)$, $\forall x\varphi(x)$, where ' φ ' may be replaced with any open formula of the language of physics, and truth-functions of those quantified statements (Quine 1960, ch. 5).

A key feature of physical theory, reflected in its logical form, is that all knowledge of physics is descriptive knowledge. The posits of physical theories are never observed directly. They are not known to us by acquaintance, with our minds reaching out to them and grasping them whole, but only as solutions to puzzles that arise about the pattern of our observations. All we know is that if our best physical theories are true, we have reason to believe that there is something or

other which answers to their quantified statements. Physical descriptions convey only general thoughts, thoughts about whatever it is that meets the criteria set out by our theory. They never convey singular thoughts about some specific individual we singled out prior to formulating a theory. As we gradually improve our theories, step by step we readjust our understanding of physical objects, too, positing, for example, sub-atomic particles instead of atoms, and space-time instead of three-dimensional space. Moving towards better explanations to account for new experimental evidence and observations, we dispense with the need for things that answer to the old descriptions, and start believing in objects that satisfy the open formulae of our new theories instead. So our grasp of physical objects is always theory-laden. We don't have theory-independent cognitive access to physical objects which outstrips the roles they play within physical theories (MacBride and Janssen-Lauret 2015, p. 297).

Every physical posit is presumed to be in the domain of physics because it satisfies a distinctive explanatory role ϕ . If our best physical theory requires for its truth that something or other answers to ϕ , we have good reason to believe in the ϕ . The explanatory roles played by our posits within physical theories impose criteria of identity on physical objects. We only have reason to believe in exactly one object per full explanatory role. A fully specified explanatory role is highly specific, including, for example, the object's exact spatiotemporal location and all its physical relations to other objects. For most physical posits, relative indiscernibility provides a full explanatory role. An exception is the case of certain quantum particles, entangled fermions, which are only weakly discernible (Saunders 2006, p. 60). Here we need an irreflexive sentence in two free variables to tell them apart. Quine attributes this relation of discriminability in two free variables to Fox, so let us call it 'Fox-discriminability' for short (Quine 1976, p. 115). Now we can conclude that all things posited by physics are subject to a generic criterion of identity: any two posits which are indiscriminable must be identified. (Here ' x and y are indiscriminable' means that either x and y are not quantum particles, and they are indiscernible, or x and y are merely weakly

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto
discernible posits like entangled fermions, and are Fox-indiscernible.)

Why do we never have good reason to posit two distinct objects x and y which are indiscernible with respect to their explanatory role in physics? Because in that case y would be completely theoretically inert. All the explanatory work that y could do is already fully taken care of by x . So we cannot justify believing in y as well as x . Our theory would carry on ticking over just as well if we had only one of them. Physical objects which satisfy exactly the same full physical descriptions (where necessary, in two variables) count as exactly one object. And physical theories regimented into canonical form cannot have two distinct but indiscernible values of variables in their domains (Janssen-Lauret 2015b, pp. 156-157).

3.2 Equivalence Relations in the Language of Physics

As a result the strongest equivalence relation for physics, as for theories regimented according to Quine's strictures generally, is slightly weaker than identity properly so called (Janssen-Lauret 2015b, pp. 155-157). Identity, in systems without constants or logically proper names, is equated with indiscernibility by the predicates of the theory. In our case, this means indiscernibility by means of the predicates of physics. Identity proper is the relation which partitions the domain into singleton equivalence classes, or simply the relation of being the same thing. This relation is stronger than any sort of descriptive indiscernibility. After all, it is always coherent to say that indiscernible objects are distinct. However detailed our physical predicates may be, it remains logically possible that x and y satisfy all and only the same open sentences (in one or two free variables) and remain distinct. We cannot derive a contradiction from conjoining the regimented version of ' x and y are indiscernible' with ' $x \neq y$ ' (Ramsey 1925, p. 31). The only exceptions are cases where one of the physical predicates is of the form ' $= x$ '. But that is ruled out in this case. The language of physics identifies physically indiscernible things, so it does not admit of a predicate which has '=' occurring within it, where '=' connotes old-fashioned sameness-of-thing

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto
identity.¹ We are only entitled to believe in objects that do some explanatory work for us, not in
theoretically inert objects that play the exact same explanatory role already performed by some
other posit (Janssen-Lauret 2015a, section 6.2).

There is always a logical possibility that two physically indiscriminable objects are distinct.
But the above result implies that physicalists can have no good reason to believe in physically
indiscriminable but distinct objects. To make the cut for being in the domain of a physicalist theory,
an object has to contribute a distinct physical explanatory role, which means that it satisfies some
open formula (possibly in two free variables) which no other object in that domain satisfies. And
beyond the use of physical predicates, there is no theoretical apparatus available to distinguish x
from y . Nor can physicalists help themselves to a stronger language, with an identity relation not
subject to the criterion of identity 'indiscriminable by the language of physics'. Physicalists must
accept the best current empirical theories used by practising scientists as true, including their
criteria of identity.

4. Impersonal vs. First-Personal Logical Form in Psychology

4.1 Impersonal Psychology?

Twentieth-century physicalists working in psychology and philosophy made explicit efforts
in a variety of ways to subject psychological theories to impersonal logical form. Traditional
behaviorists regarded first-personal methods, and data acquired by means of them, as unscientific
because they failed to resemble the impersonal methods of physics, which behaviorists regarded as
a paradigm of scientific objectivity, in one crucial way. First-personal observations are not, like the

¹ We could also derive such a contradiction in a language where the identity predicate is defined away
completely. Quine proposes such a language, retaining the old-fashioned identity predicate '=' but ditching its old-
fashioned meaning 'being the same thing' and its old-fashioned extension of all and only the domain's singleton
equivalence classes. Instead he defines the identity predicate for each language as having exactly the sense of
indiscernibility according to the predicates of that language (Quine 1961 and Quine 1970, p. 63).

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto

observations and experiments of physics, observable or repeatable independently of who the observer is. Behavior, by contrast, is intersubjectively observable whether the behavior in question is our own or other people's. It is also describable in completely impersonal terms, with no special authority being given to the first-person point of view. As a result, the project of translating the language of psychology into the language of behavior was taken to be naturalistically sound. And at the time, it really was promising from a naturalistic point of view, because behavioristic methods and theoretical statements were the order of the day among working psychologists. But psychologists and psychiatrists began to find behaviorism's strictly third-personal methods rather limiting in the 1960s and 70s (Miller 2003). First-personal methods began to be used again; nowadays they are in full swing, especially in areas like social psychology, as can be seen from my sampling of the psychology literature below. Philosophical naturalists must adjust their theories along with theoretical and methodological shifts made by practising scientists.

Post-behaviorism, many physicalists had high hopes for an eliminativist approach to psychology. Eliminative materialists predicted that psychology would fall out of use altogether and be replaced with completed neuroscience (Churchland 1981 p. 67). First-personal claims about the subject's psychological state would form no part of the evidence or theories of human psychology. They would be replaced with impersonally describable claims about brains and brain states, i.e. physical objects and events. But now, more than thirty years on, the eliminativists' predictions have not come to pass. To date there has been no sign of psychology fading away. Instead the late twentieth and early twenty-first century have seen a robust return to first-personal methods and first-personal theoretical claims. It was always a dubious move for naturalists to profess the truth of physicalism based on predictions that successors to our present best theories would be physicalist. As we reflected above, philosophical naturalists cannot rest content with vague promises for successor theories, but must accept the ontology and ideology of our present best theory. Now that thirty or more years have gone by without any clear sign of psychology being replaced by

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto
neuroscience, there is all the more reason to embrace naturalism and follow the evidence, rather than cling to physicalism in spite of the evidence.

Second philosophers who had endorsed physicalism for naturalistic reasons ought to have responded by holding on tightly to their naturalism and re-evaluating their physicalism.

Philosophers, for the most part, have done exactly the opposite of that, relaxing their definition of physicalism to 'everything is physical or supervenes upon the physical'. Introducing supervenience here loosens the tie between physicalism and the science of physics, since it requires quantification over possible worlds. Possible worlds are philosophical posits, which working physicists would either not recognize at all, or be highly dubious of. Possible-world explanations are not completely unheard of in physics. The Everett many-worlds interpretation of quantum mechanics, for instance, relies upon possible worlds. Still, such views are both highly controversial among physicists and associated with unresolved debates in the philosophy of physics as to exactly which physical entities are invoked within them (Wallace 2002). Philosophers' accounts of possible worlds generally present them as clearly abstract entities, which are not physical, but mathematical or metaphysical posits (Adams 1974, Plantinga 1974). More rarely, possible worlds are thought of, Lewis-style, as concrete objects. Even these are not uncontroversially physical entities, because some of Lewis's possible worlds contain non-physical components (Lewis 1986 p. x).

4.2 First-Personal Methods in Contemporary Psychology

Twenty-first century psychologists still use some behavioristic methods. But the use of first-personal data and statements is now also widespread. A key source of data in various branches of psychology is composed of self-reports, where subjects ascribe qualities to themselves in the first person---precisely the sort of thing behaviorists denounced as opposed to scientific objectivity.

Attitude surveys, which are very common in contemporary psychology, use such first-personal self-reports as their primary kind of evidence (Sirken et al. 1999). Grammatical differentiation between self and other is a necessary feature of this kind of methodology. Recent research into memory has

yielded results that are inexpressible without distinguishing first and third person: for instance, a case study which showed first-personal judgements to be more reliably remembered than third-personal ones by a patient with severe Alzheimer's (Klein et al. 2003). Some neuroscientists, too, are investigating the explanatory potential of the first person, and show more interest in furthering the aims of first-personal psychology than in the eliminativist physicalism beloved of several philosophers of neuroscience. In a series of experiments by Heatherton et al., for instance, brain scans revealed distinct medial prefrontal cortex activations for first-personal vs. third-personal character trait ascriptions, and the authors have drawn upon these results to engage in fruitful dialogue with social psychologists on the question of 'the self' (Heatherton et al. 2007). Active researchers in psychology have turned away from the impersonal approach of behaviorism, and the brain-state based approach of eliminativism, having found that they make better third-personal predictions and explanations of people's behavior by taking those people's first-personal statements as good though not infallible evidence about them.

In social psychology, many prominent figures go further and freely make theoretical statements about 'the self', which they take to be the referent of the first-person pronoun (Baumeister et al. 2007, Baumeister 1987, Baumeister 2010).² The self is explicitly distinguished from any 'part of the brain', 'the animal body' as well as from 'what happens inside the individual mind' (Baumeister 2010 p. 139). Characteristics ascribed to it include typically psychological ones, but also physical, interpersonal and social ones. Our cognitive access to the self, in some cases, is presumed to be essentially first-personal and introspective, a form of knowledge by acquaintance: 'one important part of the self exists mainly inside the individual's own mind ... it starts as people pay attention to themselves' (Baumeister 2010 p. 143). But we also know ourselves in more indirect ways, via knowledge by description of our own social role, or via awareness of our own bodies. For

² Special thanks to social psychologist Dr Renée Bleau for discussion of how the term 'the self' is used in social psychology, and for drawing my attention to Baumeister's work.

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto
instance, self-awareness is heightened in people who can see themselves in a mirror. Our awareness of how our bodies are perceived by others also has major repercussions for self-awareness. Eating disorders, for example, are associated with changes in self-awareness: binge eating decreases self-awareness, and is used by sufferers to escape from acute self-awareness (Heatherton et al 1991), while self-imposed dietary restraint is associated with heightened self-awareness (Bleau 1996). The logical form of the theoretical statements used by these psychologists is of a radically different kind from the impersonal form of physics: it relies essentially on grammatical distinctions between self and other, and there is good reason to believe that they cannot be translated into the impersonal language of physics. The problem is not merely that we do not have a clear translation scheme for rendering psychological claims into exactly equivalent physical claims. Although this is true, the problem runs deeper than that. First-personal theories have a richer logical form than impersonal ones, and can aspire to levels of expressive strength inaccessible to impersonal theories.

My argument in the next section combines a result of mine which uses only general principles from the philosophy of logic (Janssen-Lauret 2015a §6) with the principle of identification of physically indiscriminable objects used by philosophers of physics which we discussed in section 3 above. It will become apparent that the first-personal statements of psychology cannot be translated back into the impersonal logical form of physics without loss of information. Contemporary psychology assumes the truth of contemporary physics, but adds something extra, some expressive strength in its syntax which the language of physics lacks. In particular, the canonical language of psychology, unlike that of physics, is capable of distinguishing physically indiscriminable objects.

5. Logical Ramifications: The Increased Expressive Strength of Languages with First-Personal Logical Form

5.1 First-Personal Logical Form and Knowledge by Acquaintance

From a logical point of view, first-personal pronouns like 'I' behave differently from impersonal pronouns like 'it' or 'x'. Suppose we follow the early Russell in treating 'I' as a logical proper name (Russell 1910-11). Syntactically, it is a kind of constant, '*i*'. So where the language of physics is first-order classical logic without individual constants, the language of psychology adds to that logical form statements of the form ' φi ', with ' $\varphi i1$ ', ' $\varphi i2$ ' ' $\varphi i3$ ' for first-person pronouns with different referents. We need not prejudge the question whether the referent of the first-person pronoun is a stable self or a Humean bundle or succession of different person-stages. As the answer to this question is to some extent empirical, the logically proper names '*i1*', '*i2*' '*i3*' may be used either for different selves or persons, or for different person-stages, whichever are posited by the psychological theory at hand.

Physicalism now faces the objection that this logical form is not expressible in the language of physics. Lacking constants in their syntax, regimented physical theories have only one type of pronoun, and no analogue of 'I' or logically proper names. A canonical language with constants or logically proper names has more expressive strength than one with only impersonal pronouns like 'it', or 'x', 'y', and the like. Unlike physical theories, such a language is potentially capable of distinguishing objects which are descriptively indiscriminable. As we saw in section 3, physical posits which are physically indiscriminable are identified. The full extent of our knowledge of physical objects proceeds via physical descriptions, so where those coincide, objects must coincide.³ But this is not the case for a language which contains the additional logical form of logically proper names in its theoretical statements. Logically proper names are associated with knowledge by acquaintance rather than knowledge by description. Knowledge by acquaintance, or direct cognitive access to an object, is in itself no less naturalistic a way of knowing an object than

³ There is some debate in the literature regarding the case of bosons, which are alleged to be absolutely indiscernible (Saunders 2006, Hawley 2006). But bosons cannot be known by acquaintance, and so would not help the physicalist establish that physics does have exactly the same expressive power as first-personal psychology.

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto

indirect knowledge by description is. I have argued, for instance, that Ruth Barcan Marcus makes extensive use of knowledge by acquaintance for naturalistic reasons (Janssen-Lauret 2015b pp. 158-164). First-personal knowledge has the potential of being associated with immediate awareness of ourselves or our own current psychological states. And the psychology literature bears out that many active researchers in psychology take first-personal statements as evidence of such awareness, which is taken as a basis for self-knowledge, knowledge of the referent of the first-person pronoun. 'The first basis for selfhood is consciousness turning around on itself ... Self-knowledge would be impossible without self-awareness, which is the basic process by which attention turns around toward its source (Baumeister 2010 p. 142). Contemporary psychological evidence may include knowledge by acquaintance, reported by the subject in the first person, which bears some resemblance to what philosophers call introspective knowledge: they yield immediate knowledge of the subject's own psychological state by means of privileged access. In psychological research this is rarely used as a stand-alone method, but rather in combination with more indirect ways of knowing the self, for instance by means of knowledge of the body. For instance, attitude self-reports are more accurate when subjects are placed in front of a mirror (Pryor et al 1997).

5.2 Identity vs. Criteria of Identity

Objects known by means of acquaintance or introspection, and denoted by constants or logically proper names, resist being given criteria of identity based purely upon descriptive explanatory roles. My argument for this claim, as I noted before, is based wholly upon the philosophy of logic and some general principles about criteria of identity. It does not deploy any peculiar assumptions from the philosophy of mind or specifically anti-physicalist philosophical arguments, but flows from the difference in logical form between equivalence relation statements that deploy descriptive indiscriminability vs. identity. It is a syntactic feature of directly referential expressions, which here includes variables as well as constants, that they can legitimately be

concatenated with the identity predicate. As a result, it can be true or false *tout court* that $i1$

$= i2$, that these are the same object. By contrast, definite descriptions cannot simply be substituted for constants on either side of the identity sign; their deep logical structure is different from their surface form. 'The $F =$ the G ' is ill-formed and must be regimented to reveal its underlying logical form ' $\exists x((Fx \wedge \forall y(Fy \rightarrow x = y)) \wedge \exists z(Gz \wedge \forall w(Gw \rightarrow z = w)) \wedge x = z)$ ', with only variables on either side of the identity predicate. But as we saw in section 3, the language of physics does not have '=' in the sense of identity proper, sameness of thing; its strongest equivalence relation is indiscriminability by physical theory. So this statement would have to be weakened to 'the F is physically indiscriminable from the G '.

The difference between ' i ' and 'the F ,' and between '=' and 'is physically indiscriminable from', is not merely syntactic. It has profound philosophical consequences. Identity is the relation which relates every object to itself and to no other thing. Identity does not hold by virtue of any special condition that the object fulfils, but simply and primitively because it is the same object as itself. Now the semantic role of descriptions is to specify some condition an object must meet. An object is denoted by this description only if it meets that condition. But the semantic role of names and logically proper names is a different one: it is just to denote an object directly, without specifying anything about what the object is like. It is denoted only qua object, not qua meeting some condition. Identity, which is exactly the relation of being the same object, can only be expressed with the help of expressions which single out the object only qua object. By contrast, the fact that an object x meets some condition F a description sets out never guarantees that there is only one such object, nor that any F is the same identical x , unless identity is smuggled into the description by the use of '='. Otherwise, no matter how detailed the description, it remains logically possible that two or more things satisfy it. In the case of constants or logically proper names, by contrast, it is a crucial aspect of their semantic role that they denote that thing, and only that thing. They refer without needing to rely on descriptive conditions. If they had a descriptive component,

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto
then their denoting their referents would depend upon their meeting the condition set out in that description, as is the case for the theoretical terms of physics. And if that were so, just like for physical posits, any object which satisfied that condition would be identified with them. The strongest equivalence relation expressible by means of them would not be identity proper, quantitative identity, but some form of descriptive indiscriminability, or qualitative identity (Marcus 1961 p. 12). Constants and logically proper names, by contrast, single out only the object itself, irrespective of their qualities. So it is no coincidence, but a matter of logic, that once they have a referent, they do nothing but refer to that referent, and that they can feature in statements of quantitative identity (Janssen-Lauret 2015a § 6.2).

We saw that physical theories have a slightly weaker relation than identity as their strongest equivalence relation: physical indiscriminability. The canonical language of physics also does not have logically proper names, but only descriptions. As a result, physical theories cannot make any meaningful distinction between ' $i_1 = i_2$ ' and 'the F is physically indiscriminable from the G '. But these have distinct logical forms, and the latter by no means entails the former (Marcus 1990, p. 197). By the argument of the previous paragraph, the identity claim is stronger than the indiscriminability claim. Although every object which is physically indiscriminable from something is physically indiscriminable from itself, physical indiscriminability is logically compatible with distinctness. That means it is a weaker claim than any identity claim, since identity claims express logical facts (Janssen-Lauret 2015a, §6.2). A physicalist may object that a difference expressible only by means of expressions like ' $i_1 = i_2$ ' cannot be expressed in the language of physics, and they therefore need not take them seriously. But this just amounts to an admission that the language of psychology genuinely does have more expressive power than the language of physics. Psychologists who have embraced first-personal logical form do not deny the truths of physics, but extend physical theory with additional first-personal logical forms. Since these working scientists derive better empirical results from theories with that richer logical form, naturalists ought to

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto
embrace the logical form they introduce.

6. Conclusion: A New Challenge to Physicalism

In what sense can physicalism be maintained given these results? Physicalism based on first philosophy is dubiously motivated, but physicalists who reject first philosophy must accept the deliverances of our current best empirical theories, including psychological ones which defy formalisation in the canonical language of physics. Their theoretical statements include first-personal ones which embody a richer logical form than that of current physics. As contemporary psychology also assumes the truth of current physics, and contains a large number of statements of applied physics and biology, its canonical language provides a genuine, irreducible addition to that of physics. It adds genuinely psychological truths to the collection of physical truths. It is unclear where this leaves physicalism, the thesis that everything is physical. We derive our best and fullest knowledge of physical things from their explanatory roles described by the science of physics. Since psychology adds statements with a demonstrably richer logical form, we have good reason to avoid classifying these statements as physical knowledge. So it is far from obvious how we could have good reason nevertheless to classify them as knowledge of exclusively physical things.

A physicalist might protest that while she accepts that the methods of psychology differ from those of physics, she nevertheless maintains that their underlying subject matters coincide. One potential route to this conclusion is to say that the psychological supervenes upon the physical, but we have seen that the quantification over possible worlds that supervenience involves renders it dubious, because possible worlds are not obviously physicalistically acceptable entities. An alternative argument claims that first-personal language merely provides a new perspective upon, or mode of presentation of, familiar old physical truths.⁴ It is difficult to see how a naturalist could justify this argument. Once we admit first-personal theoretical statements, which naturalists must do

⁴ Thanks to Erasmus Mayr for this objection, even though he does not wholeheartedly endorse it.

because they are part of our current best theories of empirical psychology, and admit that these statements are not equivalent to physical statements, which we ought to because of their richer logical form, no good empirical reason remains for saying that their subject matters must coincide. To identify two posits, we need to have reason to believe they share an explanatory role. Physical posits are always introduced by means of impersonal pronouns, while some psychological posits are introduced by means of first-personal pronouns which cannot be translated into impersonal form. Our current best theories do not allow for the role of the latter to be described in the vocabulary associated with the former; so we do not have good reason to ascribe the exact same explanatory roles. My argument is not that the identification of psychological and physical posits is impossible, or that no future theory will achieve this feat. Perhaps it will. But we cannot identify x and y on the grounds that it is not impossible that they are identical. By way of comparison: it is not impossible that Julius Caesar is identical with a number, and yet we do not identify any number with the conqueror of Gaul.

A final option which is popular among twenty-first century physicalists is so-called 'non-reductive physicalism'. Non-reductive physicalists believe they are entitled to ascribe biological, psychological, mental, and social properties to objects while maintaining that the objects in question are physical. In a sense, this tallies well with the way in which social psychologists now speak about the self. Physical as well as distinctively psychological and social characteristics are freely ascribed to the self, and self-awareness and self-knowledge are taken to include knowledge of the body just as much as introspective knowledge (Baumeister 2010, p. 143). But the question now arises how we can justify calling such objects physical objects which happen to have some psychological characteristics, rather than psychological objects which have some physical characteristics. After all, we ascribe to these objects various characteristics that the science of physics makes no mention of, so it is unclear in precisely what sense they are physical. Certainly they are not readily identified with the posits of physics. Psychology builds upon the truth of

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto
physics but adds distinctively psychological truths. Physics, by contrast, remains entirely silent on
the psychological. Since psychology assumes and adds to the truths of physics, but not the other
way around, entities which answer to psychological as well as physical descriptions are more
appropriately classified as psychological entities which also have physical characteristics than as
physical entities with psychological characteristics.

In summary, recent developments in the science of psychology, especially social psychology,
appear to indicate that our best psychological theories contain first-personal theoretical statements,
with a radically different logical form from those of physics. This suggests that the forces of
physicalist imperialism, which aimed to model the methods and theoretical statements of
psychology upon physics, are now in retreat. If the physicalist is to be a good naturalist, she will
have to answer the challenge of logical form outlined here, and explain how the thesis that
everything is physical can be maintained whilst giving due diligence to first-personal theoretical
statements in the science of psychology.⁵

References

Adams, Robert Merrihew. 1974. "Theories of actuality". *Noûs*, 8(3):211–231.

Baumeister, Roy. 1987. "How the self became a problem: A psychological review of historical
research". *Journal of personality and social psychology*, 52(1):163-176.

Baumeister, Roy. 2010. "The self". In Roy Baumeister and Eli Finkel, editors, *Advanced social
psychology: The state of the science*. Oxford: Oxford University Press.

Baumeister, Roy, Kathleen Vohs, and Dianne Tice. 2007. "The strength model of self-control".

⁵ Acknowledgements: I'd like to thank Manuela Fernandez Pinto, Uskali Maki and Adrian Walsh for their efforts in putting the conference and volume together as well as for their helpful comments on a draft. Thanks are also due to the people who attended my University of Campinas seminar *Logical Form and Ontology* in the autumn of 2014, and to the audience at the international Glasgow-Campinas workshop *Quine, Science, and Ontology* where an early draft of part of this paper was presented. I am especially grateful to George Bealer, Renée Bleau, Gary Kemp, Nathan Kirkwood, Fraser MacBride, and Erasmus Mayr for discussion, and to Alan Weir for providing a response to my draft at the workshop. This research was supported by a Capes Postdoctoral Research Fellowship Grant.

Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto
Current directions in psychological science, 16(6):351–355.

Bleau, Renée. 1996. “Dietary restraint and anxiety in adolescent girls”. *British Journal of Clinical Psychology*, 35(4):573–583.

Churchland, P.M. 1981. “Eliminative materialism and the propositional attitudes”. *The Journal of Philosophy*, 78:67–90.

Duval, S., and R. A. Wicklund. A theory of objective self awareness. 1972.

Dupré, John. 1995. “Against Scientific Imperialism.” *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 2*: 374-81.

Fernández Pinto, Manuela. 2016. “Imperializing epistemology: Some shortcomings of the naturalistic turn”, this volume.

Hawley, Katherine. 2006. “Weak discernibility”. *Analysis*, 66(4):300–303.

Heatherton, T. F., and R. F. Baumeister. 1991. “Binge eating as escape from self-awareness.” *Psychological Bulletin*, 110(1):86-108.

Heatherton, T. F., A. C. Krendl, C. N. Macrae, and W. M. Kelley. 2007. “A social brain sciences approach to understanding self.” In C. Sedikides and S. J. Spencer, editors, *The Self*, 3–20.
New York: Psychology Press.

Janssen-Lauret, Frederique. 2015a. “Committing to an individual: Ontological commitment, reference and epistemology” *Synthese*. DOI 10.1007/s11229-015-0763-8.

Janssen-Lauret, Frederique. 2015b. “Meta-ontology, naturalism, and the Quine-Barcan Marcus debate”. In Frederique Janssen-Lauret and Gary Kemp, editors, *Quine and His Place in History*, 146–167. Basingstoke: Palgrave Macmillan.

Klein, S.B., L. Cosmides, and K.A. Costabile. 2003. “Preserved knowledge of self in a case of Alzheimer’s dementia”. *Social Cognition*, 21:157–165.

- Prepublication Draft, for *Scientific Imperialism*, ed. Mäki, Walsh, Fernández Pinto
Lewis, David. 1986. *Philosophical Papers*, Volume II. Oxford: Oxford University Press.
- MacBride, Fraser, and Frederique Janssen-Lauret. 2015. “Meta-ontology, epistemology and essence: On the empirical deduction of the categories”. *The Monist*, 98(3):290–302.
- Maddy, Penelope. 1997, *Naturalism in Mathematics*. Oxford: Oxford University Press.
- Marcus, Ruth Barcan. 1961. “Modalities and intensional languages”. *Synthese*, 13(4):302–322.
- Marcus, Ruth Barcan. 1990. “A backward look at Quine’s animadversions on modalities”. In R. Barrett and R. Gibson, editors, *Perspectives on Quine*, 230–243. Oxford: Blackwell.
- Miller, G.A. 2003. “The cognitive revolution: a historical perspective”. *Trends in cognitive sciences*, 7(3):141–144.
- Plantinga, Alvin. 1974. *The Nature of Necessity*. Oxford: Oxford University Press.
- Pryor, J.B., F.X. Gibbons, R.A. Wicklund, R.H. Fazio, and R. Hood. 1977. “Self-focused attention and self-report validity.” *Journal of Personality*, 45(4):513–52.
- Putnam, Hilary. 1979. “What is mathematical truth?” In *Mathematics, Matter and Method: Philosophical Papers, Volume 1*. Cambridge: Cambridge University Press.
- Quine, W.V. 1960. *Word and Object*. Cambridge, Mass.: MIT Press.
- Quine, W.V. 1961. “Reply to Professor Marcus”. *Synthese*, 13:343–347.
- Quine, W.V. 1970. *Philosophy of Logic*. Englewood Cliffs: Prentice-Hall, first edition.
- Quine, W.V. 1976. “Grades of discriminability”. *The Journal of Philosophy*, 73:113–116.
- Quine, W.V. 1981. “Things and their place in theories”. In *Theories and Things*. Cambridge, Mass.: Harvard University Press.
- Ramsey, Frank. 1931. “The foundations of mathematics”. In *The Foundations of Mathematics and other Logical Essays*. London: Kegan Paul.
- Russell, Bertrand. 1910-11. “Knowledge by acquaintance and knowledge by description”. *Proceedings of the Aristotelian Society*, 11:108–28.
- Saunders, Simon. 2006. “Are quantum particles objects?” *Analysis*, 66:52–63.

Prepublication Draft, for *Scientific Imperialism* , ed. Mäki, Walsh, Fernández Pinto

Sirken, Monroe G. , Douglas J. Herrmann, Susan Schechter, Norbert Schwarz, Judith N. Tanur, and

Roger Tourangeau. 1999. *Cognition and survey research*. New York: John Wiley and Son.

Mäki, Uskali. 2009. “Economics Imperialism: Concepts and Constraints.” *Philosophy of the Social*

Sciences 39: 351-380.

Wallace, David. 2002. ”Worlds in the Everett interpretation”. *Studies in History and Philosophy of*

Science Part B: Studies in History and Philosophy of Modern Physics, 33(4):637–661.