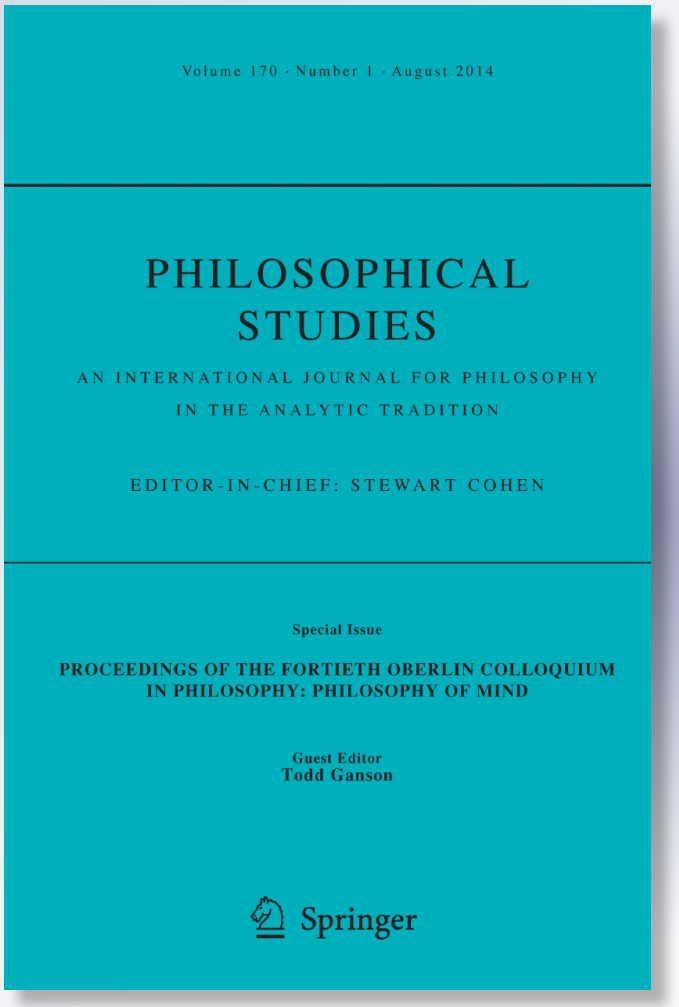
Explaining representation: a reply to Matthen

Frances Egan

Philosophical Studies
An International Journal for Philosophy in the Analytic Tradition

ISSN 0031-8116
Volume 170
Number 1

Philos Stud (2014) 170:137-142
DOI 10.1007/s11098-013-0166-y



Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Author's personal copy

Philos Stud (2014) 170:137–142
DOI 10.1007/s11098-013-0166-y

Explaining representation: a reply to Matthen

Frances Egan

Published online: 20 August 2013
© Springer Science+Business Media Dordrecht 2013

Mohan Matthen has failed to understand the position I develop and defend in "How to Think about Mental Content." No doubt some of the fault lies with my exposition, though Matthen often misconstrues passages that are clear in context. He construes clarifications and elaborations of my argument to be "concessions." Rather than dwell too much on specific misunderstandings of my explanatory project and its attendant claims, I will focus on the main points of disagreement.

1 Representationalism

My project in the paper is to argue for a particular construal of the role of representational content in computational models of cognition. The view is committed to two kinds of representational content—*mathematical* content, which characterizes the mathematical function computed by a device and subsumes both biological and artifactual computers, and *cognitive* content. The latter is determined in part by external, so-called 'naturalistic' factors (for example, visual mechanisms represent such distal properties as object boundaries, changes in illumination, and so on—properties that structure the light in relevant ways) but also by the task specification given by the theory's cognitive explanandum, as well as various pragmatic factors, as explained at length in the paper. Cognitive contents serve several important explanatory functions—demonstrating that the theory addresses the cognitive capacity that is the explanatory target of the theory, allowing us to keep track of the flow of information in the system, to name two—but they are not part of the essential individuative apparatus of the theory. In other words, the *same* computational mechanism (characterized by the function-theoretic description, as well as the algorithm and the specification of neural implementation) in a different

F. Egan (✉)
Department of Philosophy, Rutgers University, New Brunswick, NJ, USA
e-mail: fegan@rci.rutgers.edu

Author's personal copy

environment might be assigned different distal, environmentally-determined, pragmatically-motivated contents (the Visua parable in the paper spells this out).

If my construal of computational theories is correct, then insofar as one endorses the computational approach to explaining cognition one is committed to this package of representational commitments: no less (*pace* Chomsky), but also *no more*. In particular, one is not committed to the claims associated with the view that I call *hyper representationalism*—the idea that representational content is individuative of the internal states and structures posited in computational models (this is what I deny when I claim that content is not *essential*), and that such content is completely determined by a privileged, naturalistic property or relation. I am not sure if Matthen endorses *hyper representationalism*; he seems unwilling to engage the issues in the terms I have set out.

Nothing in what I have said denies that perceptual experience is representational, or that it is "phenomenologically obvious that things seem a particular way" (Matthen 2013, Sect. 1) but we should not confuse representationalism in the philosophy of perception—which holds that all phenomenological differences are representational differences—with the very general representationalist thesis endorsed by most theorists of mind, viz. that the mind is an information-using system. Explicating the phenomenology of perceptual experience is not my project.[1] Of course, we expect the sciences of the mind to eventually explain perceptual experience, and to explain how (and why) things seem the way they do. Matthen endorses the strategy that Horgan and Graham (2012) propose for the scientific explanation of perceptual experience: attribute to sub-personal, unconscious states just the contents they would have to have to link up (in one fell swoop, as it were) with what is given to us phenomenologically. It would be nice if the cognitive sciences were this easy! But whatever the prospects for this strategy, it is not (contrary to Horgan and Graham's claim) the strategy employed by computational cognitive theorists. With all due respect to commonsense, let's not confuse it with science.

On the subject of commonsense: Matthen regards as "unreasonable" my alleged "refusal to engage with [it]"(Matthen 2013, Sect. 1). Let me reiterate that the topic of my paper is computational models of cognition. But I am concerned with how *science* engages with commonsense. I point out that commonsense is the starting point for our theorizing; it specifies the explanatory target of a computational theory. I take this to be uncontroversial. But I also insist that a theory of a cognitive capacity must address our commonsense concerns; if it eschews commonsense ontologies, as scientific theories often do, then it must at some point "double back" and show that the questions motivated by commonsense have been answered. Typically, this function is served by what I call an 'explanatory gloss'—what we might think of as a 'commonsense model'—rather than by the ontology, concepts, and principles that make up the theory itself. For theories of cognition, this function is served by representational content. Recall that I criticize Chomsky for not acknowledging the importance of our commonsense interests and concerns.

[1] So I have said nothing to warrant the attributions that Matthen makes to me in his fn. 1.

Author's personal copy

Some final thoughts on representationalism: Matthen's view of computation is not sufficiently sensitive to actual practice. Not all computational models posit representations, as he suggests. Finite state machines, parallel distributed processing networks, analog relaxation systems, and massive cellular automata are all examples of computational systems with no structures that are naturally interpretable as representations. Moreover, the discussions of Putnam's understanding of Turing machines and Dennett's example of the chess-playing computer (which concerns the distinction between explicit representation and what we might call 'tacit knowledge') are orthogonal to the present issues.

2 Intrinsic content

We can talk, as Matthen does, of computational processes "extracting content", "refining content", and computationally states as "incorporate[ing]… content" (Matthen 2013, Sect. 4), if we like, but such talk needs unpacking, and unpacking this kind of talk is precisely what I have tried to do in the paper. For what it's worth, David Marr doesn't talk this way. When he said that vision "… may be thought of as a mapping from one representation to another" he wasn't talking about processes defined over contents. *Representations*, for Marr, are structures that are interpretable.

I am not sure what Matthen means by 'intrinsic content', and so I am not sure what I have been accused of in failing to recognize it (if indeed that is the charge). It is a commonplace that content, in computational models, is whatever is specified as values by the mapping from states of the device to elements of the represented domain. The idea that content is somehow intrinsic to computationally characterized states finds no support in actual computational practice. Nor does the idea of intrinsic content get any traction from either commonsense or phenomenology. As Harman (1990) and many others have noted, phenomenal experience is typically *transparent*—when we perceive a tree we do not experience any features (including representational content) as intrinsic features of the experience; we experience only features of the presented tree.

As I have said, I am not sure what the charge concerning intrinsic content is supposed to be. If the objection is to my claim that content is not essential in computational theories—that is, that environmentally-determined content is not individuative of the states, structures, processes, and mechanisms characterized by computational models—then any proposed counter-argument requires a detailed look at actual computational models and at the implications of the models for various counterfactual circumstances. This is how one establishes a claim about the individuative principles of a theory. This is not something that Matthen attempts.

Perhaps the objection is to my claim that computational theories don't posit states with *intrinsic intentionality*. If so, then I am guilty as charged. I can only reiterate my claim that computational theories are in the business of *explaining* intrinsic intentionality; they do not *posit* it. Intrinsic intentionality may be real enough, but it *is* rather mysterious—it is a legitimate *explanandum* of scientific cognitive theories; ultimately, it is not an *explanans*. My claim that intentionality may turn out to be a

Author's personal copy

feature of our experience that "do[es] not go very deep" is intended in the same vein as the following remark by Jerry Fodor:

> I suppose that sooner or later the physicists will complete the catalogue they've been compiling of the ultimate and irreducible properties of things. When they do, the likes of *spin*, *charm*, and *charge* will perhaps appear upon their list. But *aboutness* surely won't; intentionality simply doesn't go that deep (Fodor 1987, p. 97).

If intentionality doesn't appear on the final list then, assuming it really exists, it will have been reductively explained, and computational theories will likely have played an important role in the reductive explanation. My paper aims to explicate the reductive aspect of computational theories. But this possibility can only come into view once one has let go of the assumption, central to Matthen's critique, that the contentfulness of our conscious experiences requires the essential (intrinsic?) contentfulness of the subpersonal, unconscious states that give rise to them.

3 Environmental contingencies and explaining cognitive capacities

Matthen accuses me of ignoring the role of the environment in computational modeling, pointing out that

> Marr's general point was that we get great insight from modelling the visual process in a way that respects the complexities of the computational problems that perceptual systems face in the real world… these problems are often solved by reference to facts about the world in which these systems operate (Matthen 2013, Sect. 3).

Quite so, but there is no argument here for building environmental contingencies into the characterization of the system itself. Rather, Marr's point is a methodological one: the theorist must consider environmental contingencies when figuring out what the system is doing computationally. As I point out, only in some environments would computing the Laplacean of a Gaussian help an organism to see. So the computational theory must include general facts about the environment that, together with the specification of the mathematical function computed, explain the organism's success at the cognitive task. Matthen says "… environmental and task demands shape a program. If you want to know why a program has certain features, one important place to look is what it does**"** (Matthen 2013, Sect. 5). I agree, in fact, I emphasize that the *use* to which the mathematically characterized device is put in its normal environment is crucial for explaining the cognitive capacity. It is thus odd that Matthen says "According to Egan, it is a mistake to interpret the [Shadmehr/Wise] model as explaining motor control as such" (Matthen 2013, Sect. 4), especially since my discussion of the morals to be drawn from the examples begins by noting that "the Shadmehr/Wise model of motor control explains our capacity to grasp objects in our immediate environment" (Egan 2013, Sect. 5).

At times Matthen's critique seems aimed not at my view but at computationalism itself, as when he complains.

Author's personal copy

> … it is extremely problematic to assert without qualification that a neural network is computing some precise function such as a difference vector or a "Laplacean convolved with the Gaussian"… Neural networks are far too jittery for such attributions of mathematical content to be exact or exactly confirmable. For the same reasons, no biological motor system is precise enough to follow the exact navigational commands that such a computational system would issue… (Matthen 2013, Sect. 4).

It is hard to know what to make of this criticism. Computational explanations do hypothesize such functions. Yes, brains are noisy, but so are hand calculators. Any physical system computes a mathematical function only under idealization. Computational theorists are committed to the models they postulate being realized in neural structures. If it turns out that no neural structure is capable of computing the hypothesized function (under idealization, of course) then the model is simply wrong. Matthen goes on to develop his 'crucial point':

> The full system that the human motor control system employs will depend on human anatomy and human evolutionary history—on the deployment of joints and eyes and opposed thumbs like ours, and the historical contingency of what kinds of organs were present and available to be re-adapted to this task. The human system is specialized to human use; understanding the system demands local knowledge about human anatomy. This kind of understanding is not "environmentally neutral." It relies on contingencies outside the program that computes difference vectors (Matthen 2013, Sect. 4).

Again, the target seems to be not so much my view as computationalism itself. Matthen seems not to appreciate that computational theories are not in the business of answering every question one might ask about some cognitive competence. Answering so-called 'existence questions'—*why* a cognitive system is present in a particular organism—will require resources that go well beyond computational theories, though a computational theory that explains *how* an organism (with the resources available to it) solves a cognitive problem (in the environment in which it lives) will be a crucial part of the story.

Moreover, it is not obvious, as Matthen seems to assume, that representationalism is needed to answer such existence questions as "why do humans have the motor control system that they in fact have?"; otherwise, all 'existence questions' in evolutionary theory would seem to require representationalist answers. It is doubtful that representationalism will explain why some creatures have compound eyes, why some creatures are tetrachromats, and so on.

4 The role of phenomenology

I will conclude with a speculative diagnosis about what may be the deepest disagreement between the two of us. Matthen's conviction that computational theories of vision (for example) can't explain what needs explaining seems to have its roots in the idea that a genuine explanation of mental phenomena, including

Author's personal copy

perception and cognition, requires phenomenology. This view of the explanatory centrality of phenomenology may well underpin his conviction that explanatory models that don't appeal to contents that are continuous with the conscious contents of perceptual and cognitive states are completely misguided. In a word, the phenomenology of conscious experience rules, not simply in the glosses one might give to make a computational explanation perspicuous, but in the theories themselves. Computational theorizing, on this view, is deeply confused. If this is indeed the underlying worry then Matthen and I are engaged in different projects. I take the scientific practice as a given and ask how we should understand it. To imagine that we can dictate to science not only the sorts of questions it should ask but also the sorts of answers that can be given to these questions is to go well beyond the philosopher's mandate.

References

Egan, F. (2013). How to think about mental content. *Philosophical Studies*. doi:10.1007/s11098-013-0172-0.

Fodor, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.

Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives, 4*, 31–52.

Horgan, T., & Graham, G. (2012). Phenomenal intentionality and content determinacy. In R. Schantz (Ed.), *Prospects for meaning* (pp. 321–344). Boston: De Gruyter.

Matthen, M. (2013). No explanation without representation: A response to Egan. *Philosophical Studies*. doi:10.1007/s11098-013-0174-y.