Synopsis and Critical Evaluation of Robert Cummins "Representations, Targets and Attitudes"

Sitting outside at dusk one evening on a family vacation I erroneously exclaimed 'Oh a baby deer' at the sight of a rather large hare. The hare wasn't very far away from me, however her ears were partially hidden from my sight and she wasn't hopping but stalking about through the grass. While my error invited due ridicule from members of my family, most traditional theories of representation more forgivingly would have me token a |baby deer| for a 'hare'.

In Targets, Representations and Attitudes, Robert Cummins takes issue with such traditional views of representation, which claim that there is a causal relationship between states of affair in the world and corresponding internal representations that can be meaningfully manipulated in cognitive processing. To him, such views of representation (which he collectively considers 'meaning as use theories' (PAGE)) lacks explanatory power to account for cognitive error. He starts off arguing that error cannot be correctly accounted for using a theory of representation that fails to make the distinction between what a representation means (i.e. representational content, or plain 'content') what a representation intends (i.e. its intentional content or 'target'). He then goes about showing why a theory of representation needs this distinction and what ramifications this bears to related issues like attitudes such as beliefs and desires. He finally proposes a theory of mental representations grounded in mathematical isomorphism that accounts for erroroneous and proper use of representation alike.

Cummins begins by demonstrating that a meaning-as-use theory of representations is flawed. He uses the example of a chess machine engaged in endgame that falsely tokens a representation (RP3) of the position of the remaining pieces (P2) after a series of moves (M). What was needed to correctly represent position P2 was RP2, so tokening RP3 after M is error (page 6). Simply put, a representation caused by a horse "cannot be a |cow|, since it isn't caused caused by a cow. (page 5)" Likewise, the putative chess machine's error - falsely tokening RP3 *for* P2 - cannot simply be due to RP3 being falsely tokened *by* P2. If P2 needed RP2 to be correctly represented, why then was RP3 tokened? Causally mapping states of affairs to representations fails to account for error meaningfully.

Cummins goes on to show why the only plausible account for falsely tokening representations (such as RP3 for P2 in the chess machine example) must account for the discrepancy between the tokened representation and its intended use. To make that distinction he builds on two concepts: representational 'content' and the 'target' of its application, which must be kept strictly separate. "It is precisely the independence of targets from contents that makes error possible. (page 7)".

When a representation is correctly used, it must be satisfied by the state of affairs it represents during application. This state of affairs is essentially the target, while the satisfaction condition of the representation is captured by the content. This allows for a description of erroneous application of a representation as "when the target of tokening it ⌜the representation⌝ fails to satisfy its content. (page 6)". This is not a possible formulation without a clear distinction between target and content.

Furthermore, the causal independence of targets and contents must be strict. In the example of the chess machine's error, the target board configuration, P2, cannot be determined by the erroneously used representation, RP3, else there would not have been error. However, representations have the function of representing their targets, so "the notion of a representational target is essentially a functional notion, (page 7)" in other words, targets are a function of applying a representation with the intention for it to mean 'something'. For this reason, Cummins alternatively calls targets "intentional content (page 8)".

What we have now is the view that a representation's meaning is not what it contains, but what the application of its content to a particular state of affairs intends. This is is already a departure from the view that representations are causally related to what they represent. However this is not yet a satisfactory account of representation and error. Introducing the target to his theory of representation necessarily has to account for how targets are chosen. If content is independent from and cannot choose its own target, then what can?

To answer this question, Cummins posits the existence of 'intenders'. Intenders are in the business of determining what tokened content intends when it's applied to something, to quote Cummins "what ⌜the cognitive system⌝ expects to find (page 18)" during that application. This process is called 'target-fixation', since it essentially fixes content to a particular target. Aside from providing an answer to the question of how applying content to a target can happen while keeping the two independent of each other, intenders also clarify how it is possible that a representation's content can be correctly applied under some circumstances, but incorrectly applied under others. This neatly dispels the idea of 'misrepresentations' in favor of erroneous

application of representation, or erroneous target-fixation by a given intender.

This feature of intenders leads Cummins to a second departure from more traditional views of representation. So far his argument has been about the shortcomings of thinking of representations and what they represent as causally related. Now he goes on to claim that representational error cannot simply be 'misrepresentation'; Representations can only be erroneously applied. Therefore representations cannot be evaluated for whether they are true or false. While the chess machine was very much in error when it tokened RP3 to represent P2, its tokened representation was not 'false', since its content wasn't a proposition. This departure is a departure from Tarskian semantics in particular.

Cummins generalizes a larger distinction between falsehood and error from the claim that true-false distinctions don't hold for determining whether a representation is correct. "Most representations don't have propositions as contents, hence aren't the sorts of things that could be true of false anyway. (page 11)" Truth-functional evaluation of a representation then cannot be the measure of error. However a direct result of decoupling error for falsehood is that "truth isn't the opposite of error. (page 12)" How do we then determine representational error (and its opposite 'correctness')?

Cummins goes on to claim that while we cannot evaluate content for its truth-value, we can however determine its 'satisfaction condition'. In his example, RP3 represents a chess board position, hence something that cannot be propositionally evaluated, but something that has a satisfaction condition, which may or may not be me met when the the content of RP3 is applied. The focus on truth/falsehood of a representation is now shifted to the question of whether an intender can fix content on a target that meets its satisfaction condition during application.

This generalization that "error is distinct from falsehood (page 11)" has consequences beyond describing how intenders might work. Cummins cites several examples of correct mental behavior that could not possibly be accounted for without the distinction. Reductio proofs, counterfactuals, all types of thought experiments and even children playing 'let's pretend' games are among them. The crux of his argument going forward is that, rather than thinking of representations as true or false and hence correct or erroneous, we have to have a theory of representation in which content is separate from the target on which it is fixed by an intender, and in which satisfaction condition must be met when the content is applied to said target. For this, decoupling content from target and falsehood from error is crucial.

With these distinctions, especially one of a target that is functional, Cummins feels confident to tackle larger questions about attitudes, such as beliefs and desires. Cummins points out that attitudes are commonly thought of as "cognitive (or computational or functional) relations to representations, that is, as representations with a characteristic cognitive function. (page 14)" That is, a representation's attitude is endemic to that representation. Taking another look at the chess example shows why this is problematic.

When the chess machine tokens RP3 for P2 it "believes that the position after M is P2. (page 14)" However there is nothing in the machine to represents that belief directly. There is only RP3, which, as we have seen, has a satisfaction condition, i.e. that of a particular chess board configuration. Any belief about the state of affairs that representational content is meant to represent is thus not a characteristic function of that representation, therefore attitudes aren't functions of representations.

However after introducing functional notions of applications of mental content, i.e. targets, attitudes can easily be revised as a function of an application rather than a representation as well. However while targets are a function of the application of content by dint of an intender, attitudes a function of the application of content to a particular target. E.g. the application of RP3 to P2 hosts the belief attitude of the chess machine that RP3 represents P2 after M. Nothing native to RP3 itself fulfills this function. Cummins terms the application of representational content ('content') to intentional content (a 'target') "semantic content. (page 16)"

This functional notion furthermore allows for treating other error-related *malaises* of older theories of representation, notably Fodor's Representational Theory of Mind (page 15). Cummins doesn't say so explicitly, but models in which attitudes are characteristics of representations suffer from explaining what happens to representations when attitudes change due to error. A belief might change but do not cause representations to disappear. "What could it mean to retract a representation? Eliminate it from one's representational scheme? (page 14)" Erroneous applications of content to target can be retracted, however the content must somehow remain.

His discussion of error concludes with this three-way distinction of mental content: (1) representational

'content', which is dumb, hosts no beliefs about or desires of the world, (2) intentional content, a 'target' on which is a functional notion of (1) getting fixed by intenders and lastly (3) a semantic content, or 'attitude', which is a functional notion of (1) getting fixed on (2). It leads Cummins to conclude that "three distinct theories to deal with mental content" are needed, one for each type. This in mind, he develops his picture theory of representation (PTR).

PTR, while founded in part on the above distinction of three mental contents, requires another related, crucial claim: that about meaning-as-use of representation is impossible. This is founded on two observations about the role of communication and the nature of knowledge. The first is that internal representations, are not communicable material, and thus do not have meaning in virtue of its use. Cummins introduces the "distinction between meaning and *meaningfor* (page 86)" to make this point. Meaning, which his theory is trying to explain, is representation and independent of meaningfor, which is, in essence, 'what something communicates to its user(s)'. The meaningfor of something is commonly explained in virtue of its use. Internal representations however, unlike words or things that have meaningfor, "are not of interest because of what they communicate. (page 87)" Meaningfor has to be explained in terms of meaning, not vice versa.

A second, related observation is that knowledge is not the same as representation. Knowledge in PTR is, roughly speaking, schematically organized concepts, and thus closer to psychological schema-theory or "what philosophers typically call belief. (page 87)" The reason he makes the claim that knowledge is distinct from any representation is that a cognitive system may be able to exhibit knowledge of something without being able to represent it. Having concepts, organized as knowledge, can easily be misconstrued as having representations, but Cummins thinks that this claim depends on "an equivocation of two sense of 'concept', (page 89f)": one representational and one of having "cognitive grip (page 88)". The latter, which he endorses, is linked to meaning, whereas the former is related to meaningfor.

With this disclaimer, that representation cannot be explained in virtue of its use, Cummins embarks on sketching out his PTR. His first feature is that representation must be grounded in the mathematical definition of isomorphism. That is to say one thing represents another if and only if it shares the same structure, which is defined as "a set of things and a set of relations to them. (page 90)" That leads to PTR's second feature, that only structures can represent, since isomorphism is a property of two structures. Lastly, as we have seen, representations are not defined by their use, hence aren't "convention governed".

At this point it is important to note that Cummins' view of representation as isomorphic does not evolve directly and only from his tripartite theory of mental contents, rather than out of a criticism of other views of representation that are unable to conform to the three-way distinction of content, target and attitudes discussed above. He explains that isomorphism "underwrites the explanatory role of mental representation in contemporary theories of cognition. No other conception of mental representation appears to do this. (page 91)" In a way, the 'explanatory role' constraint and 'lack of alternative' are primitives in shaping his theory when answering the question 'why isomorphism?'. Rather than focussing on the lack of alternative explanations however, I will explore his argument for why isomorphism 'underwrites the explanatory role of representation'. (Note: it appears there's some debate that Millikan et al have argued the same point about isomorphism. Comparing the two standpoints is out of scope at this point.)

First off, Cummins takes on the idea that representations are mere mental 'proxies' or 'references' for the real world. He acknowledges that this idea is appealing due to its successes in other natural sciences and because it allows an appropriate reduction of phenomena otherwise too complex for finite human capacities. However it is in violation of what he calls the 'Explanatory Constraint', which underlies his criticism of all use-theories of representation. Positing things like 'proxies' or 'references' depends on a convention-governed theory of their use to "make us understand how appeals to the capacity to represent could explain cognitive capacities. (page 93)" While meaning-as-use is something Cummins eschews, he none-the-less wants to preserve some aspects of the idea that representations are like proxies. To him, the mathematical concept of isomorphism accomplishes this.

A second argument for isomorphism, which allows for preserving proxie-like features of isomorphism that aren't rule- or convention-governed, is seeing "a close analogy between representation and perception. (page 94)" For this, he cites the example of an 'autobot' navigating through a maze and thereby writing a card that is an isomorph of a car's target path. The card can then be re-used as a representation of car's target path through that maze. The car will be governed by the card, not by its target, the street/surface it runs on which

the card represents. This example shows that isomorphism grounds representation such that it "satisfies the Explanatory Constraint", "makes meaning independent of use" and "does not threaten a confusion of representational contents with targets or attitude contents (page 96)".

Cummins' writing is quite succinct and (aside from an autobot in a maze and a chess machine that somehow got to endgame only to make a rather bizarre mistake) bereft of many meaningful examples that tie his entire theory together. Perhaps it is worth revisiting my own example of error and, why, if I couldn't have had a |baby deer| representing a 'hare', Cummins has a better idea. To do so, let's rephrase my error in PTR's terms.

According to PTR my internal representation of a baby deer (previously |baby deer| which I shall call RBD), is independent of the hare (now H) which I mistook it for. RBD was fixed on H by some intender I own (perhaps several) without itself knowing anything about 'hares'. H is then a functional notion of RBD getting fixed on it by said intender(s). In addition, when I exclaimed 'oh a baby deer', I bore the attitude that I believed the hare to truly be a baby deer. RBD however contains no such information and that attitude comes from the application of RBD to H. Finally, in the overall process some comparison of H and RBD must then have taken place, to determine the accuracy of the application of RBD to H. To make the comparison, both RBD and H have to be structures of some sort, hence they are isomorphic if the application is correct. However it wasn't so H and RBD must somehow differ structurally.

This formulation of my error allows illustration of Cummins' key points: independence of content from target, independence of content from meaning, independence of meaning from use and the need for a theory of representation that makes structural comparisons. However it also allows illustration of where Cummins' theory might be wrong. Two points are noteworthy.

The first is disconcerting to most critics, as well as to Cummins himself: multiple isomorphisms. By endorsing isomorphism, PTR allows for many-to-many mappings between structures. That is my target, H, may be isomorphic with other things and my content, RBD, may be isomorphic to many other things as well. This is problematic for PTR because if weakens the special role that content and target play in cognition. It may even indirectly violate the Explanatory Constraint in that it underwrites a lot more than just "the explanatory interest of mental representation. (page 94)" It at least threatens triviality: if a hare can represent another hare in virtue of isomorphism, the specific relationship between target and content in cognition might be in danger. Cummins seems to accept this risk.

A second problem arises when one posits the functional notion of a target. In my error, H is a hare. However, it isn't. It's a structural 'something' of a hare and Cummins explicitly warns against the confusion of target (H) and referent (the hare). This is problematic for several reasons. Firstly, it requires another level of representation to enable structural comparison between target and content. Without it, isomorphism cannot work, since isomorphism requires some form of comparison. However it seems unlikely that cognitive systems have such a 'control' level of representation solely for the purpose of determining the accuracy of an application of mental content to real world objects and phenomena.

Secondly it makes the writing of new

This argument unfortunately boils down to questions of 'psychological reality' . (i.e. for RBD and H to be comparable they have to reside in the same place in the cognitive agent). However being able to compare two structures must be a given for isomorphism to work, and for that they have to, at the very least, both be accessible to some function of that agent (perhaps an intender?). Otherwise we have a notion of error that suffers from a glaring apples-and-oranges dilemma. Cummins is mute on this point.

REFERENCES

Cummins, Robert. "Targets, Representations and Attitudes", MIT Press, 1996