THEORIA

An International Journal for Theory, History and Foundations of Science

Ὁ Θεὸς ἀριθμετίζει

THEORIA

REVISTA DE TEORÍA, HISTORIA Y FUNDAMENTOS DE LA CIENCIA

AN INTERNATIONAL JOURNAL FOR THEORY, HISTORY AND FOUNDATIONS OF SCIENCE

ESTABLISHED in 1952 by MIGUEL SÁNCHEZ-MAZAS Second Series

THEORIA's website   **http://www.ehu.es/theoria**

T H E O R I A

AN INTERNATIONAL JOURNAL FOR THEORY, HISTORY AND FOUNDATIONS OF SCIENCE

CONTENTS

Still Unsuccessful: The Unsolved Problems of Success Semantics*

Javier González de Prado Salas

ABSTRACT: Success semantics is a theory of content that characterizes the truth-conditions of mental representations in terms of the success-conditions of the actions derived from them. Nanay (*Philos Stud* 165(1): 151-165, 2013) and Dokic and Engel (*Frank Ramsey* London: Routledge, 2003) have revised this theory in order to defend it from the objections that assailed its previous incarnations. I argue that both proposals have seemingly decisive flaws. More specifically, these revised versions of the theory fail to deal adequately with the open-ended possibility of unforeseen obstacles for the success of our actions. I suggest that the problem of ignored obstacles undermines success semantics quite generally, including alternative formulations such as Blackburn's.

Keywords: success semantics; naturalistic theories of content; mental representation; Frank Ramsey; teleosemantics.

RESUMEN: La Semántica del Éxito es una teoría del contenido que caracteriza las condiciones de verdad de las representaciones mentales en términos de las condiciones de éxito de las acciones que se derivan de ellas. Nanay (*Philos Stud* 165(1): 151-165, 2013) y Dokic y Engel (*Frank Ramsey* London: Routledge, 2003) han revisado esta teoría para defenderla de objeciones que socavaban sus formulaciones previas. Aquí argumento que ambas propuestas se enfrentan a dificultades decisivas. Más específicamente, estas versiones revisadas de la teoría no responden satisfactoriamente al problema planteado por la posible existencia de un número indefinido de obstáculos imprevistos para el éxito de nuestras acciones. En el artículo sugiero que la posible presencia de obstáculos ignorados supone un problema general para la Semántica del Éxito, incluyendo formulaciones alternativas como la de Blackburn.

Palabras clave: Semántica del Éxito; teorías naturalistas del contenido; representaciones mentales; Frank Ramsey; teleosemántica.

Introduction

Success semantics is a theory of mental content that accounts for the truth-conditions of beliefs in terms of the success-conditions of the actions elicited by such beliefs (Ramsey 1927; Whyte 1990, 1997; Dokic and Engel 2003, 2005; Blackburn 2010, ch. 10; Nanay 2013). Building on pragmatist themes, success semantics highlights important connections between belief and action. One reason why this project is attractive is that it seems to pave the way for a naturalistic account of mental content. In addition, it offers a plausible characterization of the content of mental representations in non-linguistic animals (Ramsey 1927; Nanay 2013).

However, despite its initial attractiveness, success semantics faces several difficulties. In particular, the possibility of unforeseen obstacles muddles the link between the success-conditions of actions and the truth-conditions of the beliefs leading to them (Brandom 1994): an action based only on true beliefs may fail due to the presence of ignored obstacles. To my knowledge, the problem of ignored obstacles has not yet been satisfactorily addressed. In this paper, I discuss this problem and critically examine recent proposed solutions. I first present an objection to Dokic and Engel's (2003) attempt to overcome the problem (section 3). Then, in section 4, I argue that the open-ended possibility of impediments for our actions puts success semantics theories in general under pressure, even if it were granted that agents somehow have representations ruling out the presence of obstacles. In the last part of the paper I claim that this issue also has problematic consequences for revised versions of success semantics that, at first sight, could seem to remain unaffected by it, in particular Blackburn's (section 5) and Nanay's (section 6).

1. Success Semantics and Ignored Obstacles

Standard versions of success semantics revolve around what Dokic and Engel call 'Ramsey's Principle' (Dokic and Engel 2003, 46):[1]

> **Ramsey's Principle (RP):** A belief's truth conditions are those that guarantee the success of an action based on that belief whatever the underlying motivating desires.[2]

The idea underwriting RP is that when a belief is true, the actions derived from it will find success. For instance, the belief that there is water in the flask will lead me to drink from it, in order to quench my thirst. If my belief is actually true, and there is water in the flask, I will succeed in quenching my thirst; however, if my belief was false and the flask was empty, it is most likely that I will fail.

---

[1] This sort of principle had been previously formulated by Whyte (1990, 150), who attributes it to Ramsey (1927, 159).

[2] I will assume that the content of beliefs can be characterized as a proposition determining a set of possible worlds (i.e. the condition in which the belief is true). This assumption is shared by the defenders of success semantics I will engage with. Due to space constraints, I leave aside the interesting question of how success semantics would fare if we characterized beliefs in a different way, for instance as a probability distribution over possible situations (when discussing Nanay's view, however, I will criticize certain probabilistic reformulations of success semantics).

RP faces several objections. In particular, contrary to what RP states, truth does not seem to be sufficient for success (Brandom 1994; Blackburn 2010; Nanay 2013).[3] One first problem is that a true belief accompanied by false collateral beliefs may lead to practical failure. The standard way of overcoming this difficulty is to say that the success-conditions of an action give us the truth-conditions of the *set of beliefs* that combine to produce that action (Whyte 1990; Dokic and Engel 2003; Brandom 1994; Blackburn 2010). One would then derive the truth-conditions of individual beliefs by contrasting the truth-conditions of all the sets in which the individual belief is involved: 'The truth conditions of the belief is the invariant of all the truth conditions of all the sets to which it belongs' (Dokic and Engel 2003, 49).

Let us grant that the problem of false collateral beliefs can be solved in this way. I will focus on a further difficulty: even when all the beliefs leading to some action are true, the action may still fail. The reason for this is that ignored obstacles or impediments may thwart the success of actions, even if such actions were based only on true beliefs (Brandom 1994; Blackburn 2010). Imagine, for instance, that I truthfully believe that I poured water in my flask; I may still fail in quenching my thirst by drinking from the flask, if I am unaware that there is a hole in it.

Thus, an action may fail without any falsity in the agent's beliefs, contrary to RP. The defender of success semantics would have to show that ignorance amounts to some sort of false belief. Perhaps it can be argued that when I intend to drink from the flask in order to quench my thirst, I have the belief that there are no holes in the flask. If this were the case, when as a matter of fact there are holes, one of my beliefs would turn out to be false, so RP would be able to account for the failure of my action. The problem is that, at least according to Brandom (1994, 177), typically there are an indefinitely large number of possible ignored obstacles for the success of our actions. Maybe the cap of the flask is stuck; or there is some substance inside that makes the water undrinkable. It seems that with a bit of imagination one will always be able to find new ways in which actions may fail due to ignored facts. I will assume that this is in general the case.

Given the assumption that there are an indefinitely large number of possible obstacles, Brandom (1994, 177) claims that it is implausible to think that agents have beliefs about each and all of them. Certainly, on the face of it such a demand does not seem realistic: it would be too cognitively burdening, at least if these beliefs are taken to be explicit or actually formed by agents (Dokic and Engel 2003, 65; also Perry 1993, 202). The problem is not only that the number of required beliefs is indefinitely large, but also that many of those beliefs will be rather bizarre and convoluted, and are not likely to have ever crossed the agent's mind (think of beliefs about the possibility of evil wizards casting defeating spells).

However, even if it is implausible to attribute to agents explicit beliefs about the absence of each possible obstacle, perhaps it is possible to ascribe to them some type of *im-*

---

[3]   Truth is not necessary for success, either. There may be lucky cases of successful actions based on false beliefs (for instance, I may succeed in catching the 4 o'clock train to London despite my falsely believing that it sets off at 5 o'clock, if at the same time my watch is one hour fast). Both Whyte (1997) and Dokic and Engel (2003, 48) claim that the truth of a belief is however necessary for the success of the *totality* of the possible actions following from it. I set this issue aside, as I will concentrate on criticizing the claim that truth is sufficient for success.

plicit no-obstacle belief. In the next section I discuss a proposal of this sort, made by Dokic and Engel (2003). After discussing the problems faced by their view, in section 4 I put forward a more general objection against the idea of rescuing success semantics by appeal to (implicit or explicit) auxiliary no-obstacles beliefs.

2. Dokic and Engel's Proposal

Dokic and Engel agree that having explicit beliefs about the absence of each possible obstacle would be too demanding. Nevertheless, they argue that such beliefs may be *implicit* in an agent who *successfully* performs an action. They work with the following notion of implicit belief:

> According to our definition, an implicit belief that *p* is a belief which has not to be considered by the agent, for instance in the form of a judgement that *p*, although the agent would be immediately justified if she were to make such a judgement on the basis of her experience. (2003, 68)

Dokic and Engel derive this notion of implicit belief from debates in epistemology. According to the epistemological view they are interested in, knowing that *p* puts the knower in a position to know that none of the possible impediments for the acquisitions of such knowledge is taking place. In this sense, the knower has implicit knowledge of the absence of all those possible impediments. So, my perceptual knowledge that it is raining puts me in a position to know that I am not dreaming, or that there is not a screen in front of me with fake rain. In accordance with this idea, Dokic and Engel propose the following principle of epistemic closure (PEC):

> **PEC**: If I know that *p*, and *q* implies that I do not, I at least implicitly know that *q* is not the case. (2003, 71)

Now, the notion of implicit knowledge proposed by Dokic and Engel is controversial, and not without problems.[4] However, let us accept it, for the sake of argument, and see now how it may be of help for success semantics.

   Dokic and Engel (2003) think that an analogous notion of implicit knowledge may be found in the case of rational agency. According to this proposal, my successful performance of an action puts me in a position to acquire knowledge that none of the impediment for the success of such an action has taken place. So, my successful action of drinking from the flask puts me in a position to know that the cap was not stuck. This does not mean that I have to form such a belief in order to be able to perform the action; rather, the idea is that, were I to form such belief, I would be justified in doing so by my experience of successfully performing the action. In this sense, Dokic and Engel would argue that, when drink-

---

   [4]   Note that agents could have this sort of implicit knowledge about things that have never crossed their minds, and that maybe they cannot even conceptualize.
        At the very least, it seems that it would be better to call this knowledge 'accessible', rather than implicit. Such knowledge would be accessible in the sense that, according to Dokic and Engel, the agent would be in a position to acquire it: it would be knowledge available to the agent.

ing from the flask, I have implicit knowledge (thereby, an implicit belief) that the cap is not stuck (and also that there are no holes in the flask, that its mouth is not blocked etc.). The relevant principle of pragmatic closure (PPC) is, according to Dokic and Engel, the following:

> **PPC**: If I am intentionally doing p, and q implies that I cannot succeed, I at least implicitly know that q is not the case. (2003, 72)

Again, let me grant that this sort of implicit belief may be attributed to agents. Even if this is so, such implicit beliefs would not rescue success semantics from the problem of ignored obstacles. The reason for this is that the sort of implicit belief Dokic and Engel characterize may only be attributed to agents in cases of *successful* performance of an action—remember that it was such success that entitled the agent to form the different particular no-obstacles beliefs.[5] However, RP will only be saved if these implicit beliefs can be attributed also in cases of *failure* due to some obstacle. The problem for RP was that when an ignored obstacle is present, the agent's action may fail even if all her beliefs were true: the truth of her beliefs would not guarantee the success of the ensuing action. So, cases of failure due to ignored obstacles will be counter-examples to RP unless it can be shown that the agent, after all, had some false belief. Given that, by assumption, all the agent's explicit beliefs are true, the defender of RP needs to argue that the failure of the action can be blamed on the falsity of an implicit belief about the absence of the obstacle. If there were such false no-obstacle beliefs, RP would correctly allow for the failure of the action—since the agent, after all, had some false belief. But Dokic and Engel cannot argue that agents have the various no-obstacle beliefs *when the action fails*, not even in an implicit way. The no-obstacle beliefs cannot be inferred from the failure of the action, so they would not be implicit—in the sense of implicit favored by Dokic and Engel—when the agent's actions fail. On the contrary, the failure of the action would at best entitle the belief that *there was some impediment* (which would be a true belief).

Dokic and Engel's proposal, therefore, does not show that agents have some false belief whenever their actions fail due to the presence of ignored obstacles. From my failure in drinking I cannot infer that there were no holes in the flask, and therefore, I cannot appeal to the falsity of such implicit belief in order to explain the failure of my action. What success semantics needs, rather than implicit knowledge, is implicit *false* beliefs, and this is not offered by Dokic and Engel's proposal.

Perhaps Dokic and Engel may reply that they only wanted to show that, when the agent has all the relevant beliefs —or at least is in a position to form them with justification—, the success of her action is guaranteed. But, what would happen then with the cases of practical failure due to ignored facts? In these cases, it would seem that there is no good reason for attributing any false belief to the agent, and nonetheless her action fails. Dokic and Engel, therefore, are just not addressing the problem of ignorance—and, thereby, they cannot defend the view that truth guarantees success.

---

[5] Dokic and Engel explicitly claim that in PPC, 'the phrase "doing p" is used to imply success' (2003, 72).

3. The Holism of Success-Conditions

I have just shown that Dokic and Engel's appeal to analogies between successful action and knowledge fails to show that agents have the sort of implicit belief discarding possible obstacles that would vindicate RP. Now I want to argue that, even if it were accepted that agents have representations ruling out each possible obstacle, success semantics would still be in difficulties. In the last part of the paper, I will claim that such difficulties burden as well the alternative versions of success semantics proposed by Blackburn (2010) and Nanay (2013).

Dokic and Engel suggest that, at least in some basic cases, agents perceive possibilities for action (i.e. affordances) in their environment, and that such perceptual representations would be false if the relevant actions failed (2003, 66-69).[6] For instance, when I reach for the glass in order to drink from it, I perceive the glass as affording my drinking from it (i.e. as affording drinkability). The accuracy of this perceptual representation would be incompatible with the presence of obstacles such as the glass being stuck to the table or its contents being too hot.

It seems that these sorts of representations could be attributed to agents independently of Dokic and Engel's proposed relation between implicit beliefs and successful actions, discussed above. One may wonder, thus, whether it is plausible to apply this idea across the board and claim that agents generally have representations (perhaps merely dispositional or implicit in some weak sense) ruling out the presence of each possible obstacle for the success of the action performed.

For the sake of argument, let us grant that this suggestion is plausible.[7] If these no-obstacle (implicit) representations are added to RP, success semantics would seem to avoid the problem of ignored impediments: RP could be reformulated so that it claimed that the success of an action is guaranteed by the truth of all the representations eliciting it (including the relevant no-obstacles implicit representations). The success-conditions of an action would determine the truth conditions of the conjunction of all these representations.

There are different ways of fleshing out these no-obstacles representations. On the one hand, there could be an indefinitely large number of specific representations, each one discarding a possible impediment. Alternatively, a single representation could rule out all the indefinitely many possible impediments. This second option was already suggested by Whyte (1990, 1997), who proposed attributing a general no-impediments belief to agents. The problems I will discuss in this section affect equally both versions of the idea, so I do not need to choose one way of developing it.

Remember that I am assuming that the number of possible impediments for an action is in general indefinitely large; if this is so, the required no-obstacle representations will

---

[6] Dokic and Engel further argue that these perceptions of affordances, when accurate, are capable of justifying an implicit instrumental belief to the effect that the agent will be successful if she performs a certain action (2003, 68). However, this claim fails prey to the objection I have presented in the previous section: such instrumental beliefs would only be justifiable —and thereby implicit in Dokic and Engel's sense— in cases of accurate perception, that is, when there are actually no obstacles. What is needed is implicit instrumental beliefs in cases when obstacles are present.

[7] A further way of motivating this possibility would be to take seriously a Bayesian picture of the mind in which the agent's priors represent the absence of obstacles and interferences.

need to rule out such an indefinitely large number of possible impediments. Under this assumption, the success-conditions of an action would give us information about the truth-conditions of a set composed by the agent's explicit beliefs *plus an indefinitely large number of no-obstacle conditions.*

The problem is that it is not clear how feasible it is to detach the truth-conditions of an individual belief from the truth-conditions of this sort of set. On the face of it, one could just contrast the success-conditions of different actions elicited by sets of representations that share the individual belief one is interested in. This is the sort of strategy typically employed in order to address the problem of detaching the contribution of collateral beliefs to the success-conditions of actions (as I discussed in section 2), and it seems promising enough when the number of collateral representations is finite and manageable. However, the prospects of this strategy are more doubtful when it is generalized to deal with no-obstacle representations, given that the number of no-obstacle conditions for each action is in principle indefinitely large. No matter how many different sets of representations (involving the individual belief in question) are contrasted, it seems that it is never guaranteed that there are no further common no-obstacle conditions shared by all such sets. Thus, it is not clear that it is possible to effectively detach the truth-conditions of an individual belief from all the no-obstacle conditions necessary for the success of the actions elicited by such a belief—especially if we take into account that in general the agent's dispositions to act on the basis of some belief will be finite and will not suffice to detach every no-obstacle condition.

Consider my belief that the water in the pot is boiling. If this belief is true, then it seems that my action of cooking fresh pasta in the pot will succeed. However, this action may fail despite such a belief being true: perhaps the pasta is rotten, preventing it from getting properly cooked. Thus, the success conditions of my action of cooking pasta in the pot are not just that the water in the pot is boiling, but also that the pasta is not rotten (and possibly further conditions). In order to factor out this no-obstacle condition, I may consider other actions derived from my belief and whose success would not be threatened by the pasta being rotten. For instance, I can use the boiling water to calibrate my thermometer to 100°C. This action would be successful even if the pasta is rotten. In turn, there seem to be obstacles for this second action that do not hinder the success of the original action of cooking pasta. Think of the presence of salt in the water. If there is enough salt, the water's boiling temperature will raise and I will fail to (accurately) calibrate the thermometer to 100°C; however, salty boiling water may still be perfectly suitable for cooking pasta.[8]

So, by contrasting the success-conditions of these two actions, one could hope to be able to filter out the contribution of the different no-obstacle conditions and in this way detach the truth-conditions of the belief that the water in the pot is boiling. Unfortunately, things are not so simple. There may always be further possible obstacles for the two actions compared and, in particular, nothing excludes the possibility that there is a common obstacle for both actions. For instance, extremely low pressure will make the temperature of the boiling water drop significantly. This would be an obstacle for both the action of cooking pasta and the action of measuring 100°C with the thermometer. Of course, there are other actions derived from the belief that the water is boiling whose success is compatible with

---

[8]   Seawater, which may be used to cook pasta, typically boils around 100.5 °C.

extremely low pressures. But there may be further common possible obstacles, more or less convoluted, far-fetched and difficult to anticipate—for instance, someone could knock over the pot one second after the pasta or the thermometer are introduced in it, frustrating both the cooking of the pasta and the calibration of the thermometer. If fanciful possibilities are allowed, it seems that we can always concoct obstacles hindering all the actions that the agent is disposed to perform on the basis of the relevant belief (but not affecting other actions) —say, a wizard could cast a spell that thwarts precisely those actions.

So, no matter how many alternative actions are considered, it is always possible that there are some further no-obstacle conditions that have not been factored out yet—and that will therefore muddle the assignation of truth-conditions to the belief we are interested in. One would never finish detaching the success-conditions associated with all the different possible obstacles for each action.[9] This problem would be especially pressing for agents with a limited repertoire of actions derivable from a given mental representation. Plausibly, young children or animals will only be disposed to use their mental representations as a guide for a reduced range of actions, which will probably be insufficient for factoring out the contribution of all the no-obstacle conditions required for the success of these actions. Note, however, that this problem does not only affect agents with a very limited set of dispositions to act. Arguably, even more sophisticated agents, like ordinary adult human beings, may lack the disposition to perform some of the actions that (given suitable goals) may follow from the relevant belief. Perhaps the agent has never considered either that action or the relevant possible obstacle. For instance, it could well be that a competent agent in the 11th century could not easily conceive of actions involving chairs in the absence of gravity. This does not mean that such an agent could not have specific beliefs about chairs. So, there is no guarantee that the *actual* behavioral dispositions of mature agents will suffice to screen out all no-obstacle conditions and thereby determine the truth conditions of individual beliefs.

Perhaps some will suggest going beyond the behavioral dispositions that the agent actually has (or would have if she entertained certain goals), and considered further indefinitely many possible actions. The problem here is that it is not clear how one should be guided when extending the agent's behavioral dispositions. Depending on how we proceed with such an extension, we will attribute different contents to the agent's beliefs. The attribution of content would remain undefined.

Thus, the open-ended possibility of obstacles introduces a far-reaching holism in the success-conditions of actions, as a result of which it does not seem possible to resort to such success-conditions in order to derive the truth-conditions of specific individual beliefs con-

---

[9] An alternative method for screening out no-obstacle conditions would be to contrast the success-conditions of an action A1 derived from the belief B1 with the success-conditions of actions A* derived from other sets of beliefs {B*} not including B1. The idea would be that the truth conditions of B1 correspond only to those parts of the success-conditions of A1 that are not shared by the success-conditions of the other actions (this method is suggested by Brandom 1994). However, this method faces similar problems to the strategy I have discussed above. First, no matter how many further actions we compare, there may always remain further no-obstacle conditions that are not shared by any of these other actions and thereby remain undetached. On top of this, this strategy will wrongly screen out as no-obstacle conditions those parts of the truth conditions of B1 that are common with the truth conditions of the beliefs {B*} involved in deriving the other actions A*.

tributing to the production of those actions. This seriously undermines success semantics, understood as the project of accounting for the truth-conditions of individual beliefs in terms of the success-conditions of the actions they produce. At best, the success-conditions of actions would allow one to determine the truth-conditions of sets of representations involving an indeterminate number of no-obstacle conditions.

To be clear, the underlying problem is not that agents will rarely be in a position to have justified beliefs ruling out far-fetched possible obstacles (for instance, obstacles associated with skeptical scenarios). The objection I am focusing on is that the actual set of behavioral dispositions associated by a competent agent to a certain belief may be insufficient to isolate the truth conditions of the belief. It is not clear that the success conditions of the actions that an agent is disposed to base on a given belief always manage to determine the specific truth conditions of such a belief, without including as well residual no-obstacle conditions. These difficulties put success semantics under a lot of pressure and they should certainly be addressed by those who want to vindicate such theories.

There remains the possibility of appealing to the success-conditions of actions in order to characterize the truth-conditions of *instrumental beliefs,* such as the belief 'That glass affords my drinking from it'. Assume that the truth-conditions of this belief are equivalent to those of the claim 'In the present circumstances, my action of drinking from the glass will succeed', a claim that would be falsified by my failure in drinking from the glass. Then, it seems that the success-conditions of my drinking from the glass will coincide with the truth-conditions of that instrumental belief. However, it would still be unclear how to get from here to the truth-conditions of ordinary, non-instrumental beliefs—such as the belief that the glass contains water, or that the water in the glass is hot. The scope of success semantics, therefore, would be rather limited.[10]

Note that RP would be trivially right for these instrumental representations, given the way in which I have defined them. However, the fact that the relation between success-conditions and truth-conditions is analytical for such representations (it is introduced by their definition) does not mean that we cannot usefully resort to the relevant success-conditions in order to identify the truth-conditions of the corresponding instrumental representation.

More generally, with the addition of a global no-obstacles representation of the form 'If B_1, B_2...B_n, then A_i will be successful', RP becomes trivially true—in particular, its truth does not depend on the specific contents of B_1, B_2...B_n or on the nature of the action A_i (Brandom 1994, 177). Ultimately, a general no-obstacles representation would just state that that the action will be successful provided that all the other beliefs it is based on are true. Once this sort of no-obstacles representation is introduced, it is automatically guaranteed that the truth of the agent's beliefs (including the no-obstacles representation) suffices for the success of her ensuing action. Brandom and Nanay take this to render RP vacuous and uninformative (Brandom 1994, 177; Nanay 2013, 154-155; also Dokic and Engel 2003, 65; Daly 2003, 60-62). However, this criticism is too quick (as already noted

---

[10] Maybe some basic, action-related representations should be characterized as instrumental, for instance as detection of Gibsonian *affordances* (which would specify opportunities for action). It can perhaps be argued that the representational capacities of some animals can be properly accounted for in terms of such instrumental representations. Note, however, that the project of success semantics is standardly conceived as aiming to deliver the truth conditions of non-instrumental, ordinary beliefs (e.g. the belief that it is raining).

by Whyte 1997, 86).[11] Even if each instance of RP becomes trivially true with the addition of the general no-obstacle premise, we can still get information about the truth conditions of some target belief by contrasting different instances of RP involving the target belief but different actions (as we did with the example of the boiling water). As we have seen, the problem for success semantics is not that it becomes vacuous, but rather that the success-conditions of the relevant actions would not allow one to detach the truth-conditions of individual, categorical (i.e. non-instrumental) beliefs from the open-ended list of no-impediment conditions.

In the next sections I critically discuss proposed solutions to the ignored obstacles problem that depart from standard success semantics. First, I address the version of success semantics defended by Blackburn (2010), who rejects the claim that truth guarantees success and focuses instead on the idea that truth is necessary for non-accidental success. Second, I examine Nanay's (2013) views, according to which truth is sufficient for raising the probability of success (even if it does not guarantee it). I will argue that both proposals are problematic.

4. Blackburn's Alternative Formulation of the Theory

The sort of holism about success-conditions that I have been discussing also makes trouble for the alternative version of success semantics defended by Blackburn (2010, 181-199), even if at first sight Blackburn's proposal appears to remain unaffected by the issue of unforeseen obstacles.

Blackburn denies that truth guarantees success. Instead, he wants to argue that truth is a necessary condition for non-accidental cases of success. The idea, roughly, is that the content of our mental representations can be characterized in terms of the conditions that explain successful episodes of acting on the basis of such representations—at least, those instances of success that are not due to accidents or sheer luck.

Blackburn's view is attractive because it seems to avoid the problem of ignored obstacles, by allowing that an action based only on true beliefs may fail if there are interfering factors. What Blackburn claims is that, when the action does succeed (not by accident), it is because the world actually was as represented by the agent. Blackburn states his proposal at the sub-sentential level,[12] characterizing the referents of the representational vehicles that compose beliefs:

> Suppose the presence of 'a' is a feature of a vehicle 'a...'. Then 'a' refers to a if and only if actual and possible actions based upon the vehicle 'a....' are typically successful, when they are, at least partly because of something about a. (Blackburn 2010, 187)

Although Blackburn's proposal is interesting, it faces several difficulties. For example, it is not clear how one may distinguish typical, non-accidental cases of success, without presup-

---

[11] Whyte's response to Brandom is in other respects misguided. In particular, he seems to conflate the attitude of believing with its content, and takes Brandom to be talking about the former when he is clearly referring to the latter (e.g. 'How could the mere existence of the beliefs which cause *A* suffice for its success?' Whyte 1997, 85).

[12] One of the reasons why Blackburn prefers to formulate his proposal at the sub-sentential level is that this makes it easier to deal with seldom (or never) used sentences (2010, 187).

posing the representational content one is trying to account for (arguably, normal success cannot just be seen as statistically prevalent cases of success). I will leave aside these potential problems and focus instead on whether Blackburn's proposal is really unaffected by the holism introduced by ignored obstacles. I want to argue that, after all, such holism does put Blackburn's version of success semantics under pressure.

As we saw above, when an action succeeds (non-accidentally), it is not only because the world actually was as represented by the agent, but also because a series of no-obstacle conditions obtained. At first sight, this does not need to cause too much trouble for Blackburn. A given representational vehicle will be typically involved in the production of several actions, and usually the success of such actions will not be associated with the same no-obstacle conditions. So, it may seem that the referent of the relevant representational vehicle can be identified as that condition which remains *invariant* across cases of (non-accidental) success for the different actions in whose production the vehicle is involved; Blackburn (2010, 188) is optimistic that this will in general be the case. The problem, however, is that some no-obstacle conditions may be shared by these different actions. There is no guarantee that, by contrasting different actions, one will eventually reach a core invariant condition that may be identified with the content of the representational vehicle in question. It could always be that some no-obstacle condition (intuitively not part of the content of the agent's representation) remains undetached, as I discussed in the previous section.

Note that these undetached no-obstacle conditions do not need to be shared by all other actions not derived from the representation in question.[13] Therefore, they cannot be factored out simply as those conditions invariantly required for the success of any possible action (i.e. for successful agency in general).[14]

One possible reply on behalf of Blackburn's theory would be to argue that the content of representational vehicles is not determined by all conditions invariant across cases of success, but only by those invariant conditions that figure in explanations of such cases of success. There may be further no-obstacle conditions that are necessary for success but that, nevertheless, are not relevant for explanatory purposes. For instance, in standard explanations of a bird's success in catching a flying insect, we do not typically appeal in an explicit way to the laws of gravitation or the atmospheric conditions that enable the bird's success.

It is not clear that Blackburn would himself favor this reply. He suggests that one needs to consider 'total explanations' of the success of actions (2010, 191); and, presumably, 'total' explanations will be exhaustive and include references to enabling, no-obstacle conditions. Be this as it may, the problem with the reply is that whether some consideration is explanatorily relevant seems to be highly dependent on the context and in particular on the interests of the individuals engaged in the explanatory activity; there does not seem to

----

[13] Imagine that all the actions performed by a bird on the basis of a representation of a flying insect depend for their success on certain atmospheric conditions. This does not mean that the success of any action possibly performed by the bird depends on such atmospheric conditions. Perhaps these conditions are not required for the success of some of the bird's actions related to other representations (say, a representation of worms under water).

[14] This strategy is suggested by some of Blackburn's remarks: 'we might wish to stress the differential or contrastive nature of explanation, in order to avoid the outcome that we are always referring to the presence of oxygen or the continuation of the gravitational field—things that are background general conditions of success' (2010, 191).

be a systematic, context-independent way of discriminating explanatorily relevant consid-
erations from facts that enable the phenomena explained but that do not need to be men-
tioned in the explanation. For instance, one may be interested in the aerodynamics of the
flight of some hunting bird and thereby appeal to details about the atmospheric conditions
when explaining the success of the bird in catching some insect. On a different occasion,
however, these atmospheric details may remain unmentioned in a nonetheless perfectly
adequate (although perhaps less complete) explanation of the bird's success. Likewise, on
some occasions one will appeal to the specific trajectory of the hunted insect in order to ex-
plain the bird's success, whereas on other occasions one may explain such success without
explicitly referring to the insect's trajectory—even if it could well be that such a trajectory
is part of what the bird represents when hunting the insect. I remain doubtful that there is
a principled way of specifying what degree of exhaustiveness should be required from expla-
nations of success determining representational content.

     I do not claim to have offered decisive arguments against Blackburn's proposal. I have
only tried to give some reasons to think that, despite appearances, the open-ended possibil-
ity of impediments for the success of our actions also ends up being problematic for Black-
burn's views. It seems that this possibility creates deep difficulties for any recognizable ver-
sion of success semantics.

     Indeed, the problems undermining Blackburn's views are arguably related to simi-
lar difficulties faced by teleosemantic theories. Teleosemantics, as presented by Millikan
(1984, 1995), can be seen as a sophisticated development of the main insights underpin-
ning Blackburn's proposal.[15] In both theories, the content of representations is given by
those conditions invariant across normal cases of successful behavior based on such rep-
resentations. The difference between the theories is that, in teleosemantics, normal con-
ditions of success are fixed by the evolutionary history of organisms—they would be the
conditions that obtained in those cases of success that explain why the representational
mechanism in question was naturally selected. As happens with Blackburn's proposal, these
normal success-conditions will typically include a large number of no-obstacles, enabling
conditions, and nothing guarantees that there will not remain some such enabling condi-
tions among the invariant success-conditions shared by all actions triggered by a given rep-
resentation (see Davies 2001 for an objection along these lines).

     Of course, assessing whether teleosemantics may overcome these difficulties requires a
much more careful discussion. However, these quick remarks suffice to show how the study
of success semantics may illuminate our understanding of the limitations of other theories
with related underlying commitments.

     In the next section I discuss how the issue of ignored obstacles undermine as well
Nanay's alternative version of success semantics.

5. Nanay's Proposal

Like Blackburn, Nanay (2013) acknowledges that truth does not guarantee success, and
he reformulates success semantics accordingly—bringing it back, he thinks, to the spirit of

---

     [15] Papineau's teleosemantics (1987), by contrast, would be closer to an evolutionary version of standard
          success semantics.

Ramsey's original views. Nanay proposes two main amendments to success semantics. First, its scope is limited to a specific class of mental representations, which he calls pragmatic representations. Second, the truth or correctness of such representations is not taken to guarantee success, but only to raise its probability.

Pragmatic representations are, according to Nanay, the immediate antecedents of action (2013, 156-159). For my purposes here, what is important is that these representations elicit and guide actions independently of other representations or beliefs of the agent. In this way, the success of the action elicited by a pragmatic representation is not affected by the possible incorrectness of surrounding beliefs and representations. By restricting the focus of success semantics to this sort of non-holistic mental representation, Nanay avoids the problem of practical failure due to incorrect collateral representations (2013, 161-162).

The problem of ignored obstacles is addressed by Nanay by endorsing a probabilistic reformulation of RP. He claims that the correctness of a pragmatic representation raises the probability of success of the action it elicits, although it is not a sufficient condition for it: there is room for failure because of unexpected impediments. For instance, even if I correctly represent the spatial location of an apple, I may fail to grab it because it is inside an invisible crystal urn, or because it explodes when touched. Nevertheless, correctly representing the location of the apple will raise the probability of my succeeding in grabbing it.

Nanay spells out the idea of probability-raising in terms of conditional probability. That the correctness of a pragmatic representation raises the probability of success of the action it antecedes means that:

> the conditional probability of the success of this action given the correctness of the representation is higher than the conditional probability of the success of this action given the incorrectness of the representation (2013, 160).

The correctness of a given representation raises the probability of success of an action *in a strong sense* if such a raise is independent of the correctness of further collateral representations, that is, 'regardless of whatever else goes on in my mind' (Nanay 2013, 160). By contrast, the representation raises the probability of success only in a weak sense if the raise depends on the correctness of collateral representations. It is the strong sense of probability raising that is relevant for Nanay's version of success semantics.

The notion of probability-raising is put to work in Nanay's specification of the content of pragmatic representations. According to Nanay,

> the correctness conditions of a pragmatic representation, R, is C if and only if C raises the probability (in the strong sense) of the success of the action R is the immediate mental antecedent of and this action is not the proper part of any other action the success of which R raises the probability of.[16] (2013, 161)

_____

[16] The final qualification in this quote has the purpose of zeroing in on the action immediately anteceded by the pragmatic representation, rather than on the series of actions that are proper parts of it; for instance, the action of moving towards an apple and touching it are proper parts of my action of grabbing it (see Nanay 2013, 161). I will be able to gloss over this subtlety in my discussion.

The problem with Nanay's proposal, I will argue, is that it fails to identify satisfactorily the content of pragmatic representations. First, there is a problem of indeterminacy: there are several different conditions that would, to some extent at least, raise in a strong sense the probability of success of the action in question. So, it is not clear which correctness-conditions should be attributed to a given pragmatic representation. Furthermore, on many occasions, the probability of success of the action would be maximally raised by a condition that is not the intuitive correctness-condition of the relevant pragmatic representation.

I will discuss these points in turn, resorting to an example in order to do so. Think of a frog catching flying targets with its tongue. In line with the sort of view about agency and mental representation endorsed by Nanay, the movement of the frog's tongue would be triggered and guided by pragmatic representations about the spatial location of the flying targets (this example should be acceptable for Nanay, since he thinks (2013, 163) that the actions of non-linguistic animals are also anteceded by pragmatic representations with correctness-conditions).[17]

Let us assume that the frog represents a target as being in a certain position S1 in its visual field and that it snaps its tongue aiming at that position.[18] One first thing to note is that the probability of success of the frog's shot is raised by several different conditions. In particular, the correctness of rather imprecise conditions would suffice to somewhat raise the probability of success. Take, for instance, the condition that the target is in the Earth's Northern hemisphere. Assuming that the frog actually is in the Northern hemisphere, the probability of success of the frog's shot is higher if the target is in the Northern hemisphere than if it were not (if the target is in the Southern hemisphere, the frog's strike is doomed to failure; if it is in the Northern, there is some chance of success).

Thus, on Nanay's view, the condition that the target is in the Northern hemisphere would seem to be a suitable candidate condition of correctness for the frog's pragmatic representation. But other less imprecise conditions of correctness would also do. Imagine that the position S1 towards which the frog actually snaps its tongue is on the right side of its visual field. Then, the correctness of the condition that the target is on the right side of the frog's visual field would raise the probability of success of the tongue's snapping (success is more probable if the target is on the right side of the frog's visual field than if it is on the left side).

---

[17] Nothing substantial in my argument hangs on the specific details of the example. Similar points can be made in relation to pragmatic representations anteceding other simple actions, such as going over a step or catching an approaching ball.

[18] Note that I am imagining a frog that represents the target being in a position S1, so it follows from the construction of the example that those are the correctness-conditions of the frog's representation. My objection will be that it is not possible for Nanay to derive such correctness-conditions from the conditions that raise the probability of success of the actions linked with that representation. I take it that it is intuitively plausible to assume that this representation would dispose the frog to do things such as shooting its tongue towards S1 (of course, the details of the connections between the frog's representations and its behavioral dispositions are a matter of scientific study). To be sure, Nanay could just claim that this assumption is wrong and that the actual relations between representations and actions are those predicted by his theory; however, as we will see, such predictions turn out to be highly counterintuitive.

Even worse, some *false* representations that are close enough to the truth also seem to have conditions of correctness that would raise the probability of success of the frog's action. The frog, remember, is aiming its shot at position S1. Consider a representation locating the target in a position S2 slightly to the right of S1. It seems that the frog's shot would be more likely to succeed on the condition that the target is in S2 than on the condition that it is not there (i.e. on the condition that the target is somewhere else in the universe). Even if the frog's shot is aimed at S1, it might deviate slightly to the right—because of some small inaccuracy in the frog's performance of the shot or because some external interference, such as a gust of wind. So, it is not too unlikely that the frog's tongue will end up in S2. Given the condition that the target is in S2, therefore, there is a fair chance that the frog's shot will succeed. By contrast, on the condition that the target is not in S2, the frog's chances of success are slim: the target could be anywhere in the universe (except in S2), so the frog is looking for a needle in a (very big) haystack.[19] Thus, the target's being in S2 would raise the probability of success of the frog's shot.

It seems, therefore, that there are an indefinitely large number of different correctness-conditions that would raise the probabilities of success of the frog's action, some of them incompatible (if the target is in S1 it cannot be in S2). It is not clear which of these conditions should be taken to constitute the representational content of the frog's pragmatic representation, according to Nanay's proposal.

One could try to say that the right correctness-conditions are the most specific ones among the possible alternatives (being in S1 would be more specific than being in the Northern hemisphere). But this will not work, since being in S2 is as specific as being in S1 and both conditions, we have seen, raise the probabilities of success.

Another option is to argue that the appropriate correctness-conditions are those that *maximally* raise the probability of success of the action. I will say that the probability of success is maximally raised by the correctness of a certain condition if the conditional probability of success given the correctness of that condition is higher than the conditional probability given the correctness of any other possible condition. This proposal, however, also gives wrong results. We have seen that the correctness of representational conditions that are *less* precise than the agent's actual representation may raise as well the relevant action's probability of success. But it can also be shown that the correctness of conditions that are *more* precise than the agent's representation may be associated with a *higher* probability raise than that yielded by the correctness of the actual representation (i.e. the representation that, given the construction of the example, the frog is assumed to have).

Imagine that the frog in our example does not represent the target as being in some specific point-like location, but rather has a more imprecise representation that places it within a certain broader spatial region S3. Imagine also that, when having such imprecise representations, the frog tends to shot its tongue towards the center of the region: it is

---

[19] Of course, on the condition that the target is not in S2 plus the further condition that it is in S1, the probability of success is very high. But this high probability would not depend exclusively on the incorrectness of the representation I am considering (i.e. a representation locating the target in S2), but also on the correctness of the further condition that the target in S1. Thus, the resulting probability raise would be weak, in Nanay's sense. Notice that the conditions of incorrectness of a representation placing the target in S2 are not restricted to the target being in S1, but also include that the target is in any other position alternative to S2.

more likely that the frog's tongue hits this central point (call it S3') than any other location within the region S3.[20] In this case, it could well be that the probability of success of the frog's shot is higher on the condition that the target is specifically in position S3' than on the more imprecise condition that the target is somewhere in region S3. Thus, the correctness of the actual representation (i.e. that the target is somewhere within region S3) does not need to maximally raise the probability of success of the action it antecedes, because such probability raise may be higher on a further, more precise, condition (i.e. that the target is specifically in location S3').

It is not only more precise correctness-conditions that can lead to a higher probability raise than that associated with the correctness of the actual representation. A higher probability raise is also produced by the correctness of conditions that rule out the presence of obstacles and external interferences. Even if the target is actually in S1, as the frog represents it to be, the frog may fail to catch it for a number of reasons. For instance, an unforeseen gust of wind may deflect the strike of its tongue; or an obstacle (say, a falling leaf) may block the frog's shot. More fancifully, the target may be protected by an invisible screen or it may explode when touched. In all these circumstances the frog's action will fail. Thus, the frog's shot will have a higher probability of success if these disturbing interferences do not take place. More precisely, the probability of success of the frog's shot will be higher on the condition that the target is in in S1 *plus* there are no gusts of wind or falling leaves blocking the way (and so on) than merely on the condition that the target is in S1—since this simpler condition is compatible with the presence of different obstacles that would thwart the success of the action. So, if we take the content of pragmatic representations to correspond to those conditions that maximally raise the probability of success, we will not be able to say that the frog's representation has as its correctness-condition just that the target is in S1; instead, we will have to ascribe to that representation a more complex correctness-condition, involving not only that the target is in S1 but also an indefinitely large list of no-obstacle conditions. Intuitively, this is not the sort of content that we would attribute to the frog's representation. Anyway, the correctness of representations with these contents would *guarantee* the success of the relevant actions (rather than merely raising its probability). RP, without amendments, would be fulfilled in relation to such representations, so it seems that resorting to probability-raising would not offer an advantage over standard success semantics.

All this seems to confirm that the correctness-conditions of pragmatic representations cannot always be identified with those conditions whose correctness maximally raises success-probabilities. It remains unclear, therefore, how Nanay may manage to select the right correctness-conditions among the different candidates that pass the test of raising the probability of success of the action anteceded by the relevant representation.

Nanay could try to reply that, in normal circumstances (that is, when interfering factors do not occur), the correctness of the agent's pragmatic representation maximally raises the probability of success of the action it triggers. This reply, however, seems problematic. First, we saw above that even in the absence of impediments and interferences, the prob-

_____

[20] Perhaps the frog starts by shooting tentatively at the center of the region where the target is (imprecisely) represented to be and, if this fails, it moves to more peripheral points of the region. The example could be constructed so that, if the target is in a non-central area of the region, it is pretty unlikely that the frog will manage to catch it.

ability raise brought by the correctness of the agent's actual representation can be lower than the probability raise associated with the correctness of a more precise condition. Second, Nanay would have to offer a principled way of specifying what circumstances count as normal (without appealing in circular ways to the content of the representations in question).[21] Moreover, if such a principled specification of normality were available, then it seems that one could just minimally modify RP by saying that truth guarantees success in normal circumstances (i.e. when interferences and impediments are absent)—Nanay's appeal to probability-raising would not be needed.

There are further worries with Nanay's proposal; for instance, it relies on substantive commitments about the representational guidance and control of actions—commitments that not everybody may want to take on board. At any rate, even if these further worries are left aside, the discussion above should suffice, I think, to make one wary of the prospects of Nanay's proposal as an account of the content of mental representations.

It is interesting to note that the problems faced by Nanay's proposal are reminiscent of analogous difficulties plaguing naturalistic informational theories of content, in particular those theories that specify the content of a representation as those conditions that show the highest probabilistic correlation with the presence of the representation (for critical discussion, see Ryder 2009). In general, there will tend to be parallels between the problems affecting informational theories (which try to extract the content of representations from the conditions that correlate with their accurate tokenings), and success semantics theories (which try to extract the content of the representation from the conditions correlated with the success of the actions derived from the representation). In both cases, there will be no-interference conditions that will hinder the detachment of the target truth conditions.

6. Conclusions

Success semantics tries to specify the content of mental representations by appeal to the connection between the truth of such representations and the success of the actions guided by them. This connection is no doubt a central aspect of intentional agency and, plausibly, will play a fundamental role in any satisfactory account of mental representation. However, I have offered reasons to remain doubtful that one can exploit this link between accuracy and success in order to derive directly the truth-conditions of individual (non-instrumental) mental representations from the success-conditions of the actions they elicit. In particular, the open-ended possibility of unforeseen disturbing interferences makes our practical success dependent on the state of the world in messy, complicated ways – which hinders (perhaps insurmountably) the formulation of an adequate theory of representation along the lines of success semantics.

_____

[21] Dokic and Engels (2003, 66) consider the possibility of characterizing normal contexts in terms of the evolutionary history of the agent (which would turn success semantics into a version of teleosemantics). The difficulties faced by this project and related ones are discussed, among others, by Hattiangadi (2007), Boghossian (1989), Loewer (1997), Ryder (2009).

Note that it will not do to define normality in terms of worlds that are similar or close enough to the actual one, because there will be cases in which the relevant interferences *actually* take place in this world.

REFERENCES

Blackburn, Simon. 2010. *Practical tortoise raising: and other philosophical essays*. Oxford: Oxford University Press.

Boghossian, Paul. 1989. The rule-following considerations. *Mind* 98: 507-549.

Brandom, Robert. 1994. Unsuccessful Semantics. *Analysis* 54 175-178.

Daly, Chris. 2003. Truth and the theory of communication. In Hallvard Lillehammer and Gonzalo Rodriguez-Pereyra, eds., *Real Metaphysics: Essays in Honour of D.H. Mellor*, 54-67 London: Routledge.

Davies, Paul Sheldon. 2001. The excesses of teleosemantics. *Canadian Journal of Philosophy* 31: 117-137.

Dokic, Jérôme and Pascal Engel. 2003. *Frank Ramsey: Truth and Success*. London: Routledge.

—. 2005. Ramsey's Principle Re-situated. In Hallvard Lillehammer and Hugh Mellor, eds., *Ramsey's legacy*, 8-21. Oxford: Oxford University Press.

Hattiangadi, Anandi. 2007. *Oughts and Thoughts: Scepticism and the Normativity of Content*. Oxford: Oxford University Press.

Loewer, Barry. 1997. A guide to naturalizing semantics. In Crispin Wright, Bob Hale and Alexander Miller, eds., Ch. 8. *A Companion to the Philosophy of Language*, Oxford: Blackwell.

Millikan, Ruth Garrett. 1984. *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press.

—. 1995. *White queen psychology and other essays for Alice*. Cambridge, MA: MIT Press.

Nanay, Bence. 2013. Success Semantics: The Sequel. *Philosophical Studies* 165: 151-165.

Papineau, David. 1987. *Reality and Representation*. Oxford: Blackwell.

Perry, John. 1993. *The Problem of the Essential Indexical, and Other Essays*. New York: Oxford University Press.

Ramsey, Frank. 1927. Facts and Propositions. *Proceedings of the Aristotelian Society* 7: 153-206.

Ryder, Dan. 2009. Problems of representation II: Naturalizing content. In Francisco Garzon and John Symons, eds., 251-279. *Routledge Companion to the Philosophy of Psychology*. London: Routledge.

Whyte, Jamie T. 1990. Success Semantics. *Analysis* 50: 149-157.

—. 1997. Success Again: replies to Brandom and Godfrey-Smith. *Analysis* 57: 84-88.

Javier González de Prado Salas obtained his PhD at the University of Southampton in 2016. His main areas of research are normativity theory, philosophy of language and epistemology.

Address: Nova Institute of Philosophy (IFILNOVA), New University of Lisbon, Faculty of Social and Human Sciences (FCSH), Av. de Berna, 26, 1069-061 Lisbon, Portugal. E-mail: jgonzalezdeprado@gmail.com