The norms of acceptance

Joëlle Proust

Institut Jean-Nicod

Abstract

An area in the theory of action that has received little attention is how mental agency and world-directed agency interact. The purpose of the present contribution is to clarify the rational conditions of such interaction, through an analysis of the central case of acceptance. There are several problems with the literature about acceptance. First, it remains unclear how a context of acceptance is to be construed. Second, the possibility of conjoining, in acceptance, an epistemic component, which is essentially mind-to-world, and a utility component, which requires a world-to-mind direction of fit, is merely posited rather than derived from the rational structure of acceptance. Finally, the norm of acceptance is generally seen as related to truth, which turns out to be inapplicable in a number of cases.

We will argue, first, that the specific context-dependence of acceptances is derived from their being mental actions, each embedded in a complex hierarchy of acceptances composing, together, a planning sequence. Second, that acceptances come in several varieties, corresponding to the specific epistemic norm(s) that constitute them. The selection of a particular norm for accepting answers to considerations of utility – to the association of an epistemic goal with an encompassing world-directed action. Once a type of acceptance is selected, however, the epistemic norm constitutive for that acceptance strictly applies. Third, we argue that context-dependence superimposes a decision criterion on the output of the initial epistemic acceptance. Strategic acceptance is regulated by instrumental norms of expected utility, which may rationally lead an agent to screen off her initial epistemic acceptance.

Introduction

An important area in the theory of action that has received little attention is how mental agency and world-directed agency interact. However one construes mental actions,[1] the analogy with world-directed action looms large. In both cases, agents have goals, preferences and reasons to act, which they can, at least in favorable conditions, articulate in practical reasoning; they can adjust their goals, and the means employed, to the ends they have, as a function of the varying constraints and opportunities. Both forms of action are often constructed as involving a form of trying, a volitional operation, which is able to execute – i.e. conduct the motor or computational program.

---

[1] For a short review, see Proust (2010)

A clear way of distinguishing a mental from a non-mental action consists in contrasting their goals, i.e. the kinds of changes that each type of action aims to bring about. While a non-mental action responds to an intention to change the external world, a mental action responds to the intention to change, or acquire, a mental property in the self. Now it can be objected, quite correctly, that any world-directed action will generate in the agent a new mental property, such as: recognizing that one's action was performed intentionally, or correctly, or in agreement with one's long-term plans and values. The point of distinguishing this kind of property from the goal of a mental action, however, is that the effort that defines a mental action as what it is (for example, an attempt to discriminate, perceptually, As from Bs) is motivated by acquiring, or making available to oneself, perceptual or memorial contents, and more generally cognitive and conative properties that would not be available without such an effortful activity. In contrast, the effort that is involved in a world-directed action is essentially aimed at producing a change in the world; it only derivatively results in the knowledge that the world has changed as intended, or in feelings and motivations of various sorts. One can capture this contrast between direct goals and derivative effects by saying that the direct goals of mental actions are reflexively cognitive or conative: they consist in changing cognitive properties in oneself; those of non-mental actions, in contrast, consist in changing the properties of one's environment, including the cognitive properties of other agents, and non-cognitive properties concerning oneself (such as health and social status).

As will be seen, however, the case of planning as a sequence of mental actions involves more than merely bringing about cognitive and conative changes in oneself: planning generally aims at acting more efficiently on the world as a consequence of one's planning. Planning thus seems to form a sui generis kind of hybrid mental and world-directed action. As will be seen shortly, acceptances are constitutive mental actions in planning. Their structure should reflect the hybrid character of world-directed planning. This, however, raises a traditional puzzle, which we can call "Jeffrey's problem". In his 1956 article, Jeffrey rejects the view that a scientist needs to accept an hypothesis in the sense of deciding "that the evidence is sufficiently strong or that the probability is sufficiently high to warrant acceptance of the hypothesis": having evidence for an hypothesis does not automatically justify acting on its basis. A standard theorem of Bayesian subjective expected utility theory is that a course of action (x_i) should be evaluated by multiplying a subjective valuation of its consequences (i.e., reward), $u(x_i)$, by their probability of occurrence $P(x_i)$. With this standard Bayesian view in mind, Jeffrey's point is that, even if a scientist came up with a single probability for an hypothesis being true, which he finds doubtful, the assignment of numerical utilities to

situations given the hypothesis would vary widely with the context of application. Therefore, even supposing that an agent has collected evidence for an hypothesis, she would not be able to accept it - i.e., act on it as if it was true. Jeffrey's problem belongs to a general class of problems that has also been explored in contextualist epistemologies: is it rational to act on a proposition on the mere basis of one's confidence in its being true? Is acceptance of a proposition rather based on its utility, i.e., on the expectations we have about the costs and benefits of acting on it? In this case, what is the status of epistemic evaluations? Do they become ancillary to interest? In order to address this question, we first need to understand the relations between propositional attitudes and the mental actions that aim to control them.

1. From passively to actively acquired attitude contents

In the traditional conception of intentional action, beliefs and desires are the basic attitudes involved in a reason to act. Restricting the prerequisites for action to beliefs and desires, however, can only offer an account of simple forms of agency, in which agents determine their goals on the basis of their passively acquired attitudes and preferences. Every student of action is familiar with the hackneyed case of the agent who desires to drink, believes that there is beer in the fridge, and therefore goes to the kitchen and opens the fridge. In higher forms of action, however, basic, passively acquired attitudes are no longer sufficient. For example, suppose an agent does not presently remember where the beer has been stored, or, although tempted to drink a beer, has a second order preference for breaking the habit of drinking beer when she is thirsty. In both cases, an epistemic or a motivational precondition of acting is not presently available; a solution, then, consists in acquiring a new epistemic or conative property. Mental actions thus allow an agent to produce in herself new beliefs or desires as a consequence of implicit or explicit self-adressed commands: epistemic: "Try to remember where the beer is stored!", conative: "Try not to let yourself want to drink beer!"

Note the difference between the epistemic and the conative forms of mental action. In an epistemic mental action, the agent does not predict what the outcome of her action will be: when trying to remember, she does not aim at acquiring the belief that "the beer is in the cellar"; she wants, rather, to know the correct answer to the question "Where is the beer stored?" Determining in advance the response to an epistemic self-query would automatically transform the mental action from an epistemic to a conative one: the agent would not want to know where the beer is, but for some reason, would want to convince herself that, say, the

beer has been forgotten at the store.[2] What would be acquired in this extreme case of self-persuasion would no longer be a belief, right or wrong, but a form of irrational acceptance.

Controlling one's mental actions has some analogies with controlling one's non-mental actions. In the latter case, the agent may need to assess beforehand whether she can perform the non-mental action (in particular when the action is unfamiliar). Subsequently, she needs to monitor how well her intended action is executed, by comparing sequentially the expected with the observed feedback until the goal is reached. In the case of mental action, an agent can similarly control her action a parte ante, by determining in advance whether her epistemic action has any chance of being successfully completed,[3] and ex post, by assessing how successful, or close to success, the action executed seems to be.[4] In both cases, the assessment of one's epistemic activities is performed by a comparator, which compares stored with observed values.[5] The mechanisms for monitoring one's knowledge and other epistemic states are key features of mental actions, as their essential function is to regulate the agents' sensitivity to norms- i.e., to allow agents to revise their mental actions when their confidence in the output obtained is below a given threshold.

The controlled vs automatic criterion helps determine the scope of mental actions. While passively believing is a truth-sensitive attitude taking perception, memory or testimony as input, judging is a mental action whose aim is to produce true beliefs as a result of an active investigation or exploration.[6] Reasoning is involved when a sequence of judgements and inferences need to be made to form an epistemic decision. Deliberating comes into play when pros and cons have to be weighted in the reasoning process.

Let us now turn to accepting, the most common in daily life of all our mental actions. It is generally recognized that acceptances, in contrast with beliefs, are voluntary.[7] Accepting, like judging, is an epistemic action, involving deliberation; when accepting P, an agent decides to

---

[2] The point of conative mental actions, in general, is to create in oneself a new motivational state, either with a given predetermined content, for example, preference for water over beer when thirsty, or with a given functional property, such as that of being a state fulfilling the requirements of a good life.

[3] This step is called "self-probing" in Proust (2008, 2010, 2012).

[4] This is the step of "post-evaluation".

[5] The feedback obtained through the epistemic comparators, however, is not directly available in perception. Recent research, reviewed in Proust (forthcoming), suggests that it is produced at a subpersonal level, through the dynamic properties of the neural activity that reliably predict outcome; the comparative chance of success of a given attempt (to remember, to solve a problem, etc.) is made available to the agent as a noetic feeling (for example a feeling of knowing, for the predictive kind, and a feeling of being correct, for the retrospective kind).

[6] There is no particular way of referring to the beliefs that result from judging, reasoning and deliberating. They can equally well be referred to as "actively acquired", as "controlled" beliefs, or as judgements.

[7] The articles of reference on accepting are: Jeffrey (1956), Stalnaker (1987), Bratman (1999), Lehrer (2000), Velleman (2000), Frankish (2004), Shah & Velleman (2005).

regard *P* as true, even though it may not be "really true".[8] As Cohen puts it, "Acceptance is a policy for reasoning, (..) the policy of taking it as a premise that *P*" (1992, 5, 7). While some authors consider that acceptances are justified when a proposition has a high probability of being true, as in the lottery paradox,[9] others deny that high probability of *P* should play any role in determining the conditions of correction for accepting *P*.[10] What justifies accepting, rather, is that "sometimes it is reasonable to accept something that one knows or believes to be false".[11] Circumstances where this is reasonable include cases where *P* "may greatly simplify an inquiry", where *P* is "close to the truth", or "as close as one needs to get for the purposes at hand". This feature of acceptance has a troublesome consequence. Due to the fact that accepted propositions are subject to contextual variation in their sensitivity to evidence and truth, they cannot be freely agglomerated in a coherence-preserving way, in contrast with beliefs.[12] A second often noted feature of accepting is that whereas beliefs and judgements are exclusively aimed at tracking the truth, acceptances seem to conjoin epistemic and practical goals. If I cannot afford to miss an appointment, I should accept, as a policy, that the bus will be late, and take an earlier one.[13]

These features of acceptance, however, fail to offer an intelligible and coherent picture of the mental action of accepting, and of its role in practical reasoning.[14] First, it is left unclear how a context of acceptance is to be construed in a way that justifies applying fluctuating epistemic standards. Is an agent who accepts propositions that she does not endorse as "really true" committed to some form of epistemological contextualism or interest-relativism?

Second, the lack of aggregativity of acceptance is a well-known source of puzzles such as the lottery and the preface paradoxes. In the lottery puzzle, an agent accepts that there is one winning ticket in the one thousand tickets actually sold. It is rational for her, however, not to accept that the single ticket she is disposed to buy is the winning one. Is this agent incoherent? In the preface puzzle, a writer may rationally accept that each statement in her book is true, while at the same time rationally accepting that her book contains at least one

---

[8] Velleman (2000), 113, Shah & Velleman (497).

[9] In the lottery paradox, it seems rational to an agent to accepts that there is one winning ticket among the thousands actually sold. But it also seems rational to her not to accept that the single ticket she is disposed to buy is the winning one. See discussion in section 3 A below.

[10] Jeffrey (1956), Kaplan, (1981), Stalnaker (1987), 92-3.

[11] Stalnaker (1987), 93.

[12] Stalnaker (1987), 92; see in particular the discussion of the preface paradox: a writer may rationally accept that each statement in his book is true, while at the same time rationally accepting that his book contains at least one error (Makinson 1965). For Stalnaker, the writer is justified in accepting both propositions, in contrast with the lottery paradox, which, according to him, does not warrant acceptance. The feature of non-agglomeration was initially introduced by Kyburg (1961), however, to account for probability-based acceptings.

[13] See Bratman (1999).

[14] For a powerful defense of this view, see Kaplan (1981).

error (Makinson 1965). Here again, is the writer incoherent? If not, why, and in which respect, is the context of her action relevant to accepting a proposition (taking it as if true)? Third, how can one possibly conjoin, in accepting *P*, on the one hand, an *epistemic requirement*, which is constitutive of the kind of acceptance it is, and, on the other hand, *utility considerations* which require an active decision as to what ought to be accepted in the circumstances?

2 - The context relevant to accepting *P*

Why is accepting *contextual*, in a way that judging, say, is not? Michael Bratman's study of planning attempts to provide an answer. Acceptances are needed as ingredients in planning. Humans need to plan their actions both because their cognitive resources and rationality are limited, and because they need to coordinate their actions with those of other agents.[15] When planning, agents need to form acceptances, as a set of context-dependent, voluntary epistemic acts, in addition to their "default cognitive background" - a set of flat-out beliefs.[16] Explaining why beliefs are not sufficient to plan one's actions is a delicate matter, however. Merely saying that acceptances, "being tied to action", are sensitive to practical reasoning is not a viable explanation: other mental actions, such as judgements, also tied to action, do not adjust their contents to accommodate considerations of practical reasoning. If acceptances are defined in terms of the utilities involved in planning (where decision strategies, such as high gain-high risk or low gain-low risk need to be made)[17], it is unclear how one can take as an *epistemic* policy that *P* is true although one believes that *not-P*. A complementary explanation by Bratman is that acceptances are context-dependent because coherence, consistency, and relevance apply within the confines of an existing plan, where the situation is modeled from the agent's viewpoint. As a result, they may rationally depart from acceptances that apply to a larger theoretical context. Presented in this way, practical acceptance again faces the problem of having to be simultaneously sensitive to two prima facie irreconcilable norms: epistemic correctness, and instrumental adequacy. How can an agent be rational in accepting a proposition *P* in spite of judging *P* to be false? If the practical context of decision is taken to directly influence the epistemic *contents* of an agent's acceptances, the epistemic mental actions, taken during planning, are taken to defer to utility in their verdicts: an unpalatable outcome for those of us who take epistemic norms to be objective requirements, indifferent to instrumental considerations.

---

[15] Bratman (1987), p. 127.
[16] Bratman (1999).
[17] As argued in Bratman (1999), 27-8.

There is, however, an alternative way in which utility determines context: not by influencing directly the epistemic contents of acceptances, but by determining the relevant norm of epistemic assessment to be applied when tentatively accepting P. This solution has been explored by Mark Kaplan (Kaplan, 1981): acceptance, according to him, does not reflect merely a state of epistemic confidence. The epistemic action involved in accepting *P* is itself driven by alternative epistemic goals and associated norms: either *exhaustivity* (tracking the "comprehensive true story about the matter", at the risk of taking false propositions to be correct, or *accuracy* (tracking the truth itself, i.e. aiming at accuracy rather than exhaustivity). Choosing one or the other strategy depends on the ends we are pursuing when we consider whether we should accept *P*. If our aim is to offer a complete and close to true picture, we are deliberately taking a risk of error: in order to avoid missing target items, we accept false positives; this is why we do not want to aggregate our acceptances. If we rather aim to be accurate, we try to produce only true statements: in order to produce only true judgements, we are ready to accept misses: now aggregation of acceptances should present no problem.[18] As a consequence of this analysis, one should index an acceptance to its relevant norm: a proposition is not merely accepted, it is rather accepted$_{at}$ or accepted$_{ct}$ (where *at* is short for: accurate truth, and *ct* for comprehensive truth). Given that the corresponding norms are different, agents should have a different assessment of their confidence when they are accepting$_{at}$ or accepting$_{ct}$ a given proposition or set of propositions. Empirical evidence shows that agents are indeed sensitive to this normative difference in their confidence judgements.[19]

3- A two-tiered view of acceptance

This article proposes a theory of acceptance that is based on the notion that utility determines a context of assessment, i.e. a specific normative angle to be used in an epistemic acceptance. Our proposal, however, differs from Kaplan's on two accounts. First, it says that there are more epistemic norms potentially involved in acceptances than the two singled out by Kaplan in the context of scientific knowledge. Second, the structure of acceptance is seen as two-tiered, with a first independently formed epistemic assessment followed by a strategic decision.

---

[18] As emphasized by Kaplan, distinguishing accuracy-driven from exhaustivity-driven acceptings allows us to deal with the preface paradox: it is rational not to aggregate one's acceptances when one's strategy is exhaustivity (one's aim, in acting, is fulfilled if one has all the relevant truths, plus some false propositions).
[19] See Koriat & Goldsmith, (1996). The claim that cognitive attitudes differ in the way they are regulated is made by Shah & Velleman (2005), p. 498. While the present chapter shares their view that "beliefs being regulated for truth is not merely a contingent fact but a conceptual truth" (500), it also claims that judgementtruth is not the only norm for acceptance.

A – The variety of the epistemic norms of assessment determines types of acceptance

As we saw above, a proposition can aim at truth under a norm of strict accuracy, or of comprehensiveness. Accepting as epistemically certain or as uncertain is another case of attitude toward truth, mediated by an informational parameter that restricts the domain of evaluation for this acceptance to a given set of worlds. As shown by Yalcin (2007), the semantics of epistemic modal judgements involves interestingly non-standard truth-conditions. Other forms of acceptance, however, are not aiming at truth (even though they may in certain conditions be truth-conducive). Given the need for epistemic coordination with other agents, acceptances can be driven by a norm of consensus; the condition of correction for accepting$_{con}$ *P* is that the agent is disposed to take *P* as a fact because the other agents in her reference group do. When consensus works as an epistemic norm for an acceptance, the existence of a common disposition among agents in the reference group is not a contingent fact resulting from a shared common background of flat-out beliefs.[20] The agents, rather, deliberately form their acceptances as a function of those of others (e.g., while defending a client in court, planning peace talks, or conducting organizational communication). Once a policy, a set of consensual beliefs, or an overall plan is accepted$_{con}$, however, one needs to monitor the coherence of one's current acceptances with former ones and with background beliefs, filtering out those that do not match. Coherence is thus the driving norm for an additional type of acceptance. For example a novelist needs to monitor the coherence of her factual descriptions: she needs to accept$_{coh}$ every situation she imagines as her novel evolves. Thus norms for accepting *P* can be used either in isolation or in combination with another. Two types of acceptance are of particular interest in the context of verbal communication. Acceptance-for-intelligibility aims to recognize the meaning of a proposition as easily accessible to a recipient. Easily understood sentences require less effort from the hearer. When planning a speech or a written communication, or when hearing or reading one, a trade-off between intelligibility and amount of detail, among other factors, is necessary. A proposition that is accepted$_{int}$ is one that meets the standard for communicating information in an efficient way.[21] A related norm for accepting a given set of sentences is relevance. Here, the agent needs to deal with a trade-off between the informativeness of a message – the added inferential means it provides to reason about a situation – and the additional resources

---

[20] A case of non-normative consensuality is exemplified in Koriat (2008), where common acceptances result from a similar epistemic background and similar apparent fluency, and not from any attempt to accept the same propositions as others do.
[21] For an epistemologist's defense of the value of understanding, see Kvanvig (2005).

required to process this message.[22] Agents need to appreciate and accept (or reject) a proposition under a norm of relevance (accept$_{rel}$) in order to understand/produce messages with their intended inferential potential. Such norms not only regulate communication: they play a constitutive role in the organization of plans, which must also find a balance between level of detail and ease of memorization.

In sum: there are various epistemic norms constituting what it is to accept P, generating different types of conditions of correctness, i.e. different semantic rules, and, at the regulation level, different types of confidence judgements in agents. Recognizing this diversity offers a natural way out of the puzzles related to aggregating beliefs, mentioned above. Concerning the preface puzzle: if the author's epistemic goal is one of offering an ideally comprehensive presentation of her subject matter, it will not be contradictory for her to accept$_{ct}$ all the sentences in her book, while accepting$_{pl}$ (accepting as plausible or likely) that one of them is false. Hence, a mental act of acceptance$_{ct}$ does not allow aggregation of truth, because its aim is exhaustive (include all the relevant truths) rather than accurate truth (include only truths). Similarly, in the lottery puzzle, an agent may accepts$_{at}$ that there is one winning ticket in the one thousand tickets actually sold, while not accepting accepting$_{pl}$ that the single ticket she is disposed to buy is the winning one.

It is important to appreciate that, although the selection of a particular epistemic goal responds to the practical features of one's plan, there is no compromise between epistemic and instrumental norms concerning the *content* of acceptances. Agents' epistemic confidence in accepting$_n$ (i.e., accepting under a given norm n), is not influenced by the cost or benefit associated with being wrong or right.[23] For example, one may need to aim at retrieving an accurate, or alternatively an exhaustive list of items from one's memory; both types of aim correspond to different epistemic goals, different correctness conditions, and will generate different judgements of confidence in the epistemic content so produced. Which type of acceptance is selected depends on its instrumental role within a plan. But epistemic acceptances are normatively autonomous: they respond to the standard that constitutes them, and to nothing else. Thus we do not endorse the view that an epistemic decision to accept P entails yielding to utility considerations.[24] In the view defended, utility drives the selection of

---

[22] Sperber & Wilson (1986/1995). Speaking of a trade-off entails that the agents have the ability to compare the acceptances reached under the two conflicting norms.

[23] Costs and benefits are here meant to refer to those incurred in acting on one's acceptances (including reporting them). Whether there can be purely epistemic costs is discussed in Joyce (1998).

[24] For an extensive discussion of the autonomy of epistemic requirement relative to instrumental considerations, see Broome (1999), Proust (2012).

a specific norm; epistemic content, however, is in itself indifferent to utility. And it should be. Why? Because utility may vary unexpectedly within a single action, and still more so across time, when planning precedes execution by several weeks or months. Had an agent not formed her acceptance independently of utility, she would have no informational map of the situation on which to base her decision of how to act, in a given context and in a given epistemic state relative to that context. This independence is a condition for rationality. Changing the stakes can affect how we act on the world, not how we think about it.

B – From epistemic to strategic acceptance

The reader may correctly object, at this point, that the theory of epistemic acceptance developed above, although recognizing that the selection of an epistemic norm depends upon utility, is still failing to address the function of acceptances in practical reasoning. We now need to understand how the *output* of an epistemic acceptance so construed is adjusted to the final ends of the plan. Should an epistemic content, accepted with a degree of confidence *c*, be used in action when the stakes are high? [25] If this question makes sense, the decision to act on one's epistemic acceptance, - strategic acceptance- constitutes a second step in accepting *P*. Utility does not merely influence the selection of certain epistemic norms of acceptance – such as accuracy, exhaustivity or consensus. It also influences decision to act in a way that may depart more or less from the cognitive output of epistemic acceptance. For the type of context that now gains currency has to do with maximizing the expectation of good rather than with the context of selecting a given norm, and reaching the associated epistemic evaluation. Let us suppose, for example, that an agent has opted for a strategy of exhaustivity in trying to retrieve information (e.g. in trying to reconstruct a shopping list). Let us assume that she currently accepts with confidence c= 70%, that her reconstructed list is exhaustive. She now needs to estimate whether this confidence level is sufficient to act on, given an expected ratio between benefit and cost, say, of 3 in the present context.[26] The chances of this action being successful can be assessed on the basis of its table of utilities: the agent is in a position to decide whether, given the interests involved, world uncertainty and her own

---

[25] The fact that an epistemologist like Mark Kaplan is mainly interested in the acquisition of scientific knowledge may explain why he concentrates only on the question of how the norms of acceptance are influenced by epistemic utility. The latter kind of utility, however, does not exhaust utility: as Richard Jeffrey has shown, knowledge is meant to influence action in the world, which tries to maximize the expectation of good in a number of ways (Jeffrey, 1956, 245). Invoking the relationship between knowledge and action has fueled the intuition that knowledge ascription depends on the context of its use in action (Stanley, 2005). On a construal where the selection of the relevant epistemic norm is context-sensitive, this attractive idea is not incompatible with the present view of the autonomy of epistemic normativity, as will be seen below.

[26] For various rational strategies of decision, see Jeffrey, (1956).

performance assessment, she should use her epistemic acceptance or not, opt for a new policy (say, continue searching, or ask someone) or not.

Some readers, however, might be tempted to reject the intellectualist step altogether, and consider that strategic acceptance is all there is to accepting; on this view, epistemic acceptance – conducted with no pragmatic interest in mind – is just an idealization of traditional epistemology: it plays no role in decision. A conceptual and an empirical argument can be levelled against this suggestion. The conceptual argument was already offered above in section 3B. Given that utility varies across time, agents must have a way to determine what they are ready to accept from an epistemic viewpoint, and how confident they are in accepting it, independently of how they can strategically use this acceptance. The existence of an autonomous level of epistemic acceptance, as was argued at the end of section A, allows agents to have a stable epistemic map that is independent from local instrumental considerations. An empirical argument in favour of a two-step theory of acceptance is that the strategic step can, in fact, be dissociated from the epistemic step. There are contexts of planning and acting where an agent has no strategic leeway: in "a forced-choice task", a subject is not offered the possibility of selecting the kind of acceptance$_n$ she wishes to perform: the task dictates which norm is relevant, and the option of deciding not to act on this acceptance is not left open to the subject. Take the case of a multiple choice questionnaire where students need to identify correct algebraic identities: the task is accuracy-driven and does not include an option of not-responding. Neither is strategic acceptance an option when the agent has no access to a table of utilities for a given context of action. In this case, epistemic acceptance will be the only step guiding planning and action.

In contrast, when participants in a memory experiment are allowed to freely volunteer or withhold their answer, i.e., when strategic acceptance is open to them, they can substantially enhance the accuracy of their report compared to a situation where they are forced to respond. What happens in the free-report case is that subjects can refrain from reporting their memory of an item when their confidence is moderate, but also when there is a high penalty for giving an incorrect answer. On the basis of their experimental data in metamemory, Koriat & Goldsmith (1996) are able to conclude that strategic regulation, i.e. a decision to volunteer or to withhold an epistemic response, involves three mechanisms, two of which correspond to our two types of acceptings:

1. A monitoring mechanism is used to assess the correctness of a potential epistemic response (probability of being correct): this is our epistemic accepting, a mental action that terminates with a confidence judgement of a given level.

2. A decision mechanism is used to compare the probability of being correct as assessed in (1) and a preset response criterion probability, whose threshold is set on the basis of implicit or explicit payoffs for this particular decision (this is our strategic acceptance).

3. A control mechanism must finally take action in accordance with what is strategically accepted.[27]

Being conditional on variation in utility, the strategic step of acceptance becomes particularly cogent in contexts where subjective prediction is made difficult by environmental or internal variance, and where there is a significant difference between the costs and benefits associated with a given decision to act based on acceptance or rejection of proposition *P*.

We began our discussion of acceptance with the problem of having epistemic standards fluctuate with contexts, which constitutes a serious threat for the rationality of practical reasoning. Now we see that, on the two-step view, there is no such fluctuation. Note that the only rationally promising option an agent has for strategically controlling her previous epistemic acceptance consists in *screening out* answers that fall below her threshold of subjective confidence given her decision criterion. She does not have the option of enhancing the overall correctness of her acceptance, unless, of course, she is given a second chance to form a better-informed acceptance. Thus rational deliberation, in planning an action, does not lead agents to make irrational bets on how the world is, beyond what they feel they know; it rather presses them to use their knowledge cautiously, in a context-sensitive way.

There are, however, pathological cases – addiction, phobia, schizophrenia, brain lesions - where agents, lacking "control sensitivity", decide what to do independently of their own epistemic acceptance, and of its specific confidence level.[28] The existence of a selective deficit in rational decision suggests that epistemic and strategic acceptances are cognitively distinct steps.

In summary, this section has argued that in situations where an agent can freely consider how to plan her action, knowing its stakes or assessing them probabilistically, she can refrain from acting on what she has epistemically accepted. When no such option is offered to her, however, an agent acts exclusively on the basis of her epistemic acceptance. This two-step theory accounts nicely for the cases of acceptances discussed in the literature. Judging *P* true flat-out is an accepting under a stringent norm of accurate truth, while "judging *P* likely" is an accepting under a norm of plausibility, conducted on the background of probabilistic beliefs

---

[27] Koriat & Goldsmith, (1996), p. 493. See also Goldsmith & Koriat (2008), p. 9.
[28] On such dissociation in schizophrenia, see Koren et al. (2006).

regarding *P*. Adopting *P* as a matter of policy divides into accepting, under a norm of consensus, a set of premises to be used in collective reasoning, and accepting under a norm of coherence, (as in judgements by contradiction, legal reasoning, etc.). Assuming, imagining, supposing do not automatically qualify as acceptances. Only their controlled epistemic forms do, in which case they can be identified as forms of premising. The preface and the lottery paradoxes, unpalatable consequences of classical acceptance, are dissolved once the appropriate distinctions between types of acceptance, and associated semantics, are made

Our theory predicts that errors in acceptances can be either instrumental, epistemic or strategic. Instrumental errors occur when selecting an epistemic norm inappropriate to a context (e.g., trying to reconstruct accurately a forgotten shopping list, when comprehensiveness is sufficient). Epistemic errors can occur either in misapplying a selected norm to a given cognitive content (for example, seeming to remember accurately that *P* when *P* is merely imagined); or in forming an incorrect judgement of confidence about one's epistemic performance (e.g., being highly confident in having correctly learned an item in a list when one will actually fail to retrieve it). Appropriate confidence judgements have a crucial epistemic role as they filter out a large proportion of first-order epistemic mistakes. Strategic errors, finally, occur when incorrectly setting the decision criterion given the stakes (e.g., taking an epistemic acceptance to be non-important in its consequences on action when it objectively is). Some potential objections, however, need to be briefly examined.

4. Objections and replies

A – Acceptance does not form a natural kind

It might be objected that, if acceptance can be governed by epistemic norms as disparate as intelligibility, coherence, consensus and accuracy, it should not be treated as a natural kind. To address this objection, one needs to emphasize that normative diversity in acceptances has become salient in metacognitive studies, where agents were seen to opt for accuracy or exhaustivity, or to use fluency as a quick, although loose way, of assessing truthfulness.[29] Normative diversity results from the fact that agents have different ways of capitalizing on informational states, and that different regulative requirements correspond to them. What makes accepting a unitary mental action is its particular function: that of adjusting to various standards of utility the epistemic activity associated with planning and acting on the world. This adjustment requires both selecting the most promising epistemic goal, and suppressing those acceptances that do not meet the decision criterion relevant to the action considered.

---

[29] Koriat and Goldsmith (1996), Reber and Schwarz (1999).

B – Sophistication implausible

A second objection might find it implausible that ordinary agents have the required sophistication to manage acceptances as described, by selecting the kind of epistemic acceptance that is most profitable given a context of planning, by keeping track of the implicit or explicit payoffs for a particular option, and by setting on this basis their response criterion probability. It must be acknowledged that agents do not have, in general, the conceptual resources that would allow them to identify the epistemic norm relevant to a particular context. Acceptances, however, can be performed under a given norm without this norm being represented explicitly. Agents learn implicitly that a given norm governs acceptances performed in a given task and context: such learning is apparent from the way in which agents practically monitor their acceptance, i.e. express confidence levels reliably correlated with a given norm (such as accuracy, comprehensiveness or coherence). Agents thus rarely need to deliberate about the kind of accepting appropriate to a context, because the selection is often dictated by the task or triggered by the motivation for an outcome: At the supermarket counter, the exact change is expected. When doing math, an accurate answer. At the bus stop, an approximate waiting time. At a family meeting, a consensual conception of a situation. In this variety of contexts, no reflection is needed: agents are trained, by prior feedback, to select the proper acceptance.[30] In certain circumstances, it is to be expected that conflicts of acceptances will occur. The conflict accuracy-comprehensiveness, discussed earlier, arises in memorial tasks as well as in the context of scientific inquiry. In religious cognition, epistemic authority and consensus-based acceptances may be overriden by considerations of intelligibility, coherence or plausibility. These conflicts, again, can be solved without having to explicitly identify the epistemic norms underlying the respective forms of acceptance. A change in context and in the associated motivations points to the kind of acceptance that should be preferred. The implicit character of the selection of a given type of acceptance is incompatible with the view that personal-level prior intentions are necessary to cause mental actions. It is argued in Proust (2012) that a mental action usually results from the realization that one of the epistemic preconditions for a developing embedding world-directed action is not met.

Now the problem of over-sophistication can also be raised about strategic acceptance: agents clearly do not perform explicit statistical calculations about expected performance and distance from a criterion value. A short answer is, again, that they do it implicitly, in a fairly

---

[30] Velleman takes acceptance to be a subdoxastic attitude (Velleman, 2000, 246). On the view defended here, mental actions, including acceptances, could not be properly monitored if they were entirely subdoxastic. What is suggested, rather, is that mental actions can be selected implicitly through context-generated motivations

reliable way. There is no consensus, at present, nor even a complete theory, about how agents manage to integrate cognitive information about the probability of predicted consequences for each option with the associated reward motivations and risk aversion in a single quick and timely decision. Concerning decision making, however, robust evidence indicates that the ability to re-experience an emotion from the recall of an appropriate emotional event is crucial in integrating the various values involved in an option (Gibbard 1990, Bechara, Damasio and Damasio 2000). Agents are guided in their strategic acceptance by dedicated emotions (with their associated somatic markers), just as they are guided in their epistemic acceptance by dedicated noetic feelings. (Koriat 2000, Hookway 2003, Proust 2007). The probabilistic information about priors, on the other hand, seems to be automatically collected at a subpersonal level (Fahlman, Hinton and Sejnowski 1983).

C – Value pluralism and epistemological relativism

Third, some epistemologists might observe that such a variety of epistemic standards paves the way for epistemic value pluralism, i.e., the denial that truth is the only valuable goal to pursue. Our variety of epistemic acceptings should indeed be welcome by epistemic pluralists, who claim that coherence or intelligibility, are epistemic goods for their own sake.[31] It is open to epistemic value monists, however, to interpret these various acceptances as instrumental steps toward acceptance$_{at}$, i.e. as "epistemic desiderata", in the terms of Alston (2005). Let us add, however, that, in contrast with the epistemological project of studying what constitutes knowledge or success in inquiry, the present project aims to explore the multiplicity of acceptances open to lay persons, given the informational needs that arise in connection with their daily ends.

A further worry is that recognizing that the selection of acceptances is guided by instrumental considerations may seem to invite a relativist view about epistemic norms. Epistemic relativism is the view that what constitutes epistemic success (in particular, knowledge that P) depends on the standards used in a context of assessment. "Standards" here refers not directly to the practical import of accepting P, but to the level of certainty, or evidentiality, that is required to attribute knowledge to the agent.[32] For an epistemic relativist, it can be true to say that "Joe knows that his car is parked in his driveway" (P_1) in a low-standard context, and that "Joe does not know that his car is parked in his driveway" (P_2) in a high-standard context. This is so, from a relativist viewpoint, because variable contexts of assessment (i.e. variable standards) determine variable knowledge attributions. The question,

---

[31] See Kvanvig (2005), DePaul (2001).
[32] See McFarlane (2005).

then, is whether our various epistemic acceptances are based on various contextually-driven standards. Given our view that norm selection depends on the ends pursued, are not our acceptances also standard-relative?

Let us emphasize, first, that selecting an acceptance determines a context of *assessment* in a different sense from that of McFarlane (2005): as was argued above, epistemic assessment of a given acceptance depends on the particular norm that guides it; McFarlane's assessment, on the other hand, exclusively concerns knowledge attribution. Second, on the present view, low and high standards are used to assess not epistemic, but strategic acceptance. Let us assume that what changes when one accepts P_1 or P_2 above is not determined by the way the world is, but by the utility of accepting one or the other.[33] Then, on the present view, the confidence one has in P_1 and P_2 should, rationally, stay invariant across contexts of epistemic assessment. What should vary is one's willingness to act on it, i.e. strategic assessment. It would thus seem natural, from our perspective, to interpret McFarlane's concept of knowledge as an acceptance assessed as "true enough given the stakes" (a strategic acceptance), rather than as an acceptance assessed as "true under a norm of accuracy" (an epistemic acceptance). McFarlane, however, does not claim that low/high standards refer to utility; he rather sees them as epistemic requirements (associated, for example, with the kind of scrutiny involved in a sceptical argument versus an easygoing ordinary attibution). From the present viewpoint, these various epistemic requirements determine different forms of acceptance (such as accepting$_{at}$ vs. accepting$_{pl}$), which can be respectively assessed in an objective way. Whether these various types of acceptance equally deserve to be called "knowledge" is another matter, that we cannot discuss here. What can be concluded is that our notion of acceptance is meant to keep epistemic evaluation separate from the strategic decision to use it in action. It does not embody, as such, a relativistic view about epistemic norms.

Conclusion

The purpose of the present contribution has been to clarify the norms respectively involved in mental and world-directed action, through an analysis of the case of the mental action of acceptance. This type of action is relevant to our problem because it is both an epistemic type of mental action, sensitive to multiple norms such as truth and coherence, and a major constituent in planning world-directed actions, sensitive to considerations of utility. Dick Jeffrey found acceptance problematic because it did not seem rational to act on a proposition

---

[33] This assumption is needed in order to distinguish a contextualist from a relativist view about knowledge attribution.

on the mere basis of one's confidence in its being true. Our two-tiered theory of acceptance proposes an answer to Jeffrey's worry. It is argued that acceptance needs to include two distinct sequential steps: epistemic acceptance and strategic acceptance. Instrumental considerations, however, are appealed to in selecting a particular epistemic norm for accepting *P*. Multiplicity of acceptances is a consequence of bounded rationality: Given their limited cognitive resources, agents need to focus on the specific epistemic goals likely to offer the best return in epistemic correctness and practical utility. According to context, they may aim at accuracy, comprehensiveness, plausibility, intelligibility, coherence, or consensus. Even though utility influences the selection of a type of epistemic acceptance, it does not influence its epistemic output – neither in its content, nor in the degree of confidence related to it. Strategic acceptance, however, can screen off given epistemic acceptances that do not reach a decision criterion. This two-tiered conception fulfills the requirements of using cognitive resources to further one's ends without dissolving epistemic into instrumental norms, or ignoring the practical demands that world-directed actions address to active thinking.

References

Alston, W. (2005). *Beyond "Justification": Dimensions of Epistemic Evaluation*, Ithaca, NJ: Cornell University Press.

Bechara, A. Damasio, H. and Damasio, A.R. (2000). Emotion, Decision Making and the orbitofrontal cortex. *Cerebral Cortex,* 10: 295-307.

Bratman, M.E. (1987). *Intention, Plans, and Practical Reason*, Cambridge: Harvard University Press.

Bratman, M.E. (1999). *Faces of intention*. Cambridge: Cambridge University Press.

Broome, J. 1999. Normative requirements, *Ratio* 12, 398-419.

Cohen, L. J. (1992). *An Essay on Belief and Acceptance.* Oxford: Clarendon Press.

DePaul, M. (2001). Value Monism in Epistemology. In Steup, M. (ed.), *Knowledge, Truth, and Duty. Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford: Oxford University Press, 170-185.

Fahlman, S.E., Hinton, G.E. and Sejnowski, T.J. (1983). Massively parallel architectures for A.I.: Netl, Thistle, and Boltzmann machines. *Proceedings of the National Conference on Artificial Intelligence*, Washington DC: 109-113.

Frankish, K. (2004). *Mind and supermind*. Cambridge: Cambridge University Press.

Gibbard, A. (1990). *Wise Choices, Apt Feelings*. Oxford: Clarendon Press.

Goldsmith, M. and Koriat, A. (2008). The strategic regulation of memory accuracy and informativeness. *The Psychology of Learning and Motivation*, 48, 1-60.

Hookway, C. (2003). Affective States and Epistemic Immediacy. *Metaphilosophy* 34: 78-96.

Jeffrey, R. C. (1956). Valuation and acceptance of scientific hypotheses. *Philosophy of Science,* 23, 3, 237-246.

Joyce, J.M. (1998). A Nonpragmatic vindication of Probabilism. *Philosophy of Science*, 65, 4, 575-603.Kaplan, M. (1981) Rational acceptance. *Philosophical Studies*, 129-145.

Koren, D., Seidmann, L.J., Goldsmith, M. & Harvey P.D. (2006) Real-World Cognitive—and Metacognitive—Dysfunction in Schizophrenia: A New Approach for Measuring (and Remediating) More ''Right Stuff'', *Schizophrenia Bulletin, 32, 2, 310-326.*

Koriat, A. (2000). The Feeling of Knowing: some metatheoretical Implications for Consciousness and Control. *Consciousness and Cognition*, 9, 149-171.

Koriat, A. (2008). Subjective confidence in one's answers: The consensuality principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 34(4): 945-959.

Koriat, A. & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review,* 103, 3, 490-517.

Kvanvig, J. (2005). Truth is not the primary epistemic goal, in M. Steups and E. Sosa, (eds). *Contemporary Debates in Epistemology*, Oxford: Blackwell, 285-296.

Kyburg, H.E. (1961). *Probability and the Logic of Rational Belief,* Middletown, CT: Wesleyan University Press.

Lehrer, K. (2000). Discursive Knowledge. *Philosophy and Phenomenological Research*, 60, 3, 637-653.

Makinson, D. C. (1965). Paradox of the Preface, *Analysis,* 25: 205-207.

McFarlane, J. (2005). The assessment sensitivity of knowledge attributions. In T.S. Gendler and J. O'Leary-Hawthorne (eds.), *Oxford Studies in Epistemology*, vol.1, Oxford: Clarendon Press, 197-234.

Proust, J. (2007). Metacognition and metarepresentation : is a self-directed theory of mind a precondition for metacognition ? *Synthese*, 2: 271-295.

Proust, J. (2010). Mental acts. In: T. O'Connor & C. Sandis (eds.) *A Companion to the Philosophy of Action.* Wiley-Blackwell, 2010, 209-217.

Proust, J. (2012). Mental acts as natural kinds. In A. Clark, J. Kiverstein & T. Vierkant (eds*), Decomposing the will*, Oxford: Oxford University Press.

Reber, R. and Schwarz, N. (1999). Effects of perceptual fluency on judgements of truth. *Consciousness and Cognition*, 8: 338-342.

Shah, N. & Velleman, J.D. (2005). Doxastic Deliberation, *The Philosophical Review,* Vol.114, 4, 497-534.

Sperber, D. & D. Wilson. (1986/1995). *Relevance: Communication and Cognition.* Oxford: Blackwell.

Stalnaker, R. (1987). *Inquiry.* Cambridge: MIT Press.

Velleman, J.D. (2000). *The possibility of practical reason.* Ann Arbor: The University of Michigan Library.

Yalcin, S. (2007). Epistemic Modals. *Mind*, vol. 116, 464, 983-1026.