



LUND UNIVERSITY

Goal Rationality in Science and Technology: An Epistemological Perspective

Olsson, Erik J

Published in:

How technology shapes science: philosophical perspectives on the role of technology in science

2014

[Link to publication](#)

Citation for published version (APA):

Olsson, E. J. (2014). Goal Rationality in Science and Technology: An Epistemological Perspective. In S. O. Hansson (Ed.), *How technology shapes science: philosophical perspectives on the role of technology in science* Springer.

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Goal rationality in science and technology: an epistemological perspective

Erik J. Olsson

Olsson, E. J. (2014). Goal Rationality in Science and Technology: An Epistemological Perspective. In Hansson, S. O. (Ed.) *How technology shapes science: philosophical perspectives on the role of technology in science*. Springer.

Abstract: According to one strong intuition, what distinguish science from technology are the ultimate goals of these activities: while the goal of technology is practical usefulness, the goal of science is truth. The question raised in this paper is whether, and to what extent, this means that goal setting rationality is also different in the two domains. It is argued, preliminarily, that it is not: the theory of goal rationality in management and technology can be profitably transferred to the scientific context. This conjecture is substantiated partly by remarking on its intrinsic plausibility and partly, and above all, by appealing to its systematic advantages. As for the latter, the conjecture is applied to four closely related epistemological debates, with pragmatist ingredients, concerning truth as a goal of inquiry. It is argued that these otherwise puzzling debates can, in this way, be fruitfully reconstructed and perhaps even resolved.

Keywords: goal rationality, technology, truth, knowledge, Rorty

1. Introduction

Authors reflecting on the distinction between science and technology often share the intuition that what demarcates the one from the other are their ultimate goals: while the goal of technology is practical usefulness, the goal of science is truth. For instance, Jarvie (1972) concludes that “[t]echnology aims to be effective rather than true” (p. 55), and Skolmowski (1972) that “[s]cience concerns itself with what is, technology with what is to be” (p. 44). The same basic intuition is clearly stated, more recently, in Houkes (2009):

The intuition is that technology is, in all its aspects, aimed at practical usefulness. Thus whether technological knowledge concerns artefacts, processes or other items, whether it is produced by engineers, less socially distinguished designers, or by

consumers, the prima facie reason to call such knowledge “technological” lies in its relation to human goals and actions. And just as scientific knowledge is aimed at, or more tenuously related to, the truth, so technological knowledge is shaped by its relation to practical usefulness. (p. 312)

A few paragraphs later, Houkes summarizes this intuition – the TU-intuition as he calls it – in the following terms:

[T]he TU-intuition understands the difference between natural science and technology (or, more narrowly, the engineering sciences) in terms of a difference in goals: the former aims at finding out true theories, where the latter aims at practical usefulness. (p. 318).

Despite its many adherents, the TU-intuition has been challenged from various camps. The intuition presupposes that the kind of truth at which scientific activity is directed is not reducible to practical utility. Some instrumentalists may want to question that presupposition in favor of a “pragmatic” theory of scientific truth. From such a standpoint, the distinction between technology and science as a distinction between goals becomes difficult to uphold. Another possibility is to accept the realist concept of truth and to accept also that scientific theories are candidates for this property but to deny that scientists may justifiably accept or reject a theory because of its truth-likeness. It could be maintained, instead, that theory choice ought to be governed by the usefulness of theories for solving empirical and theoretical problems of science. If this is correct, then the proposal that technology and science differ regarding their goals becomes, once more, problematic.¹

Yet it should be made clear that these challenges to the TU-intuition are based on rather extreme views about the notion of truth that arguably relatively few authors are willing to subscribe to, at least in unqualified terms. Moreover, the TU-intuition does not strictly speaking presuppose, in a strong sense, a realist or objectivist concept of truth. What it does assume is, again, that the kind of truth that is relevant for scientific theories is not reducible to mere practical utility. For instance, a theory which rejects any notion of scientific truth beyond empirical adequacy would still be sufficient for the purposes of underwriting the TU-intuition (so long as empirical adequacy is not itself reduced to practical utility).

¹ For a review of various attempts to undermine the TU-intuition the reader is referred to the excellent exposition in Houkes (2009).

Although my later discussions will shed doubt on some pragmatist proposals, I shall in the following basically take the TU-intuition for granted. My main question will rather be this: given that science and technology differ with respect to the goals that they aim at, does that also mean that they differ regarding goal *rationality*? Or is the rationality involved in setting technological goals basically the same as the rationality involved in setting scientific goals? Goal rationality has been studied extensively in management science, and some of that work has been transferred to a technological context. This has given rise to a fairly precise and well-developed framework within which technological goal rationality can be fruitfully studied. The study of goal rationality in science, by contrast, turns out to be a surprisingly underdeveloped intellectual territory.

The hypothesis to be substantiated in this paper states that the theory of goal rationality in management and technology can be profitably transferred to the scientific domain. Hence, even if the TU-intuition is correct, so that there is a fundamental difference in the ultimate goals that science and technology aim to attain, the qualitative difference between these two phenomena should not be overestimated: they still share the same goal rationality. I will argue in favor of this conjecture partly by remarking on its intrinsic plausibility, but also partly, and above all, by appealing to its systematic advantages. Concerning the latter, I will suggest that the conjecture can be fruitfully employed in reconstructing and, at least to some extent, resolving four related epistemological debates concerning the truth as a goal of inquiry. Two of these debates are pragmatist in spirit, thus presenting potential threats to the TU-intuition along the lines suggested above. I will provide reasons for thinking that these approaches rely on an incomplete understanding of the nature of goal rationality.

2. Goal rationality in technology

Goal rationality has been studied extensively in management theory, where it is central in so-called MBO, an acronym standing for Management By Objectives (e.g. Mali, 1972). This has led to the development of a common approach, codified in the acronym SMART, according to which goals should be Specific, Measurable, Achievable, Realistic and Time-bound. This theory has been refined and systematized by Sven Ove Hansson and his research group at the Royal Institute of Technology (KTH) in Stockholm (e.g. Edvardsson and Hansson, 2005). In the following, I will refer to the framework developed by Hansson et al as SMART+, signaling that it represents an updated, philosophically more sophisticated, version of the original SMART conditions. (This is my terminology, not theirs.) The KTH group has used

the theory in its study of environmental objective (Edvardsson, 2004) and transport objectives (Rosencrantz et al, 2007). I will discuss its relevance for technology below.²

A goal is typically set for the purpose of achieving it. We will say that a goal is *achievement-inducing* if setting it furthers the desired end-state to which the goal refers. Thus the goal of becoming rich is achievement-inducing (for me) if my setting that goal makes it more likely that I will in fact become rich, e.g. by inspiring me to focus on accumulating wealth, which may eventually lead to my actually becoming wealthy. As a first approximation, a goal G is achievement-inducing for a subject S just in case the probability that S attains the goal G is increased by S setting herself the goal G, i.e., in semi-formal terms, just in case $P(S \text{ attains the goal } G \mid S \text{ sets herself the goal } G) > P(S \text{ attains the goal } G)$.

Edvardsson and Hansson proceed to use the notion of achievement-inducement to define the concept of goal *rationality*: in their view, a goal is rational if it performs its achievement-inducing function (sufficiently) well. This is a satisficing rather than an optimizing notion of rationality (Simon, 1956). Evidently, in order to be achievement-inducing and therefore, on this proposal, rational a goal should guide as well as motivate action. One could also argue that rational goals serve to coordinate actions among several agents, but that aspect will not play any major role in the following.

There is certainly more to be said about this proposed concept of goal rationality. First, as it stands it begs the question against visionary goals such as “world peace” or, in general, goals that cannot be fully attained. An example from Swedish transport policy is the so-called “vision zero” goal stating that, in the longer run, no one should be killed or seriously injured as the effect of a traffic accident (Rosencrantz, Edvardsson and Hansson, 2007). A goal that cannot be attained is not achievement-inducing and hence irrational according to the proposed definition. However, there is an obvious way to avoid this untoward result by redefining achievement inducement. A goal G is achievement-inducing for a subject G, on the revised proposal, just in case the probability that S attains the goal at least partially or, alternatively, at least approaches the attainment of G, is increased by S setting herself the goal G.

Second, achievement-inducement, even in the less demanding sense, cannot be all there is to goal rationality. If it were, the rational thing to do would be to set oneself trivial goals that can be easily attained: poking one’s nose, lifting one’s hand, and so on. The likelihood that I

² The account of SMART+ in this section draws mainly on Edvardsson and Hansson (2005). The reader is advised to consult that paper for additional references.

manage to raise my hand if I set myself the goal to do so is very close to one. Goals which are more difficult to achieve, such as getting oneself a solid education, would be dismissed as irrational. However, the proposal does make good sense as a tie-breaking condition in a setting where there are already a number of candidate goals that have been singled out on the basis of other considerations. Faced with a set of goals that are equally attractive in other respects, it is reasonable to select one that is achievement-inducing.

With these clarificatory remarks in mind, what does it mean, more specifically, that a goal can guide and motivate action? It is useful at this point to distinguish between three types of criteria of goal-rationality: those related to what the agents *know*, what they *can do* and what they *want to do*. From the first, epistemic perspective, goals should be *precise* and *evaluable*. A goal such as “achieving a better society” fails on the first account, that of precision. That goal is not very useful for guiding action unless supplemented with more precise instructions. There are at least two different aspects of precision: directional and temporal. A goal is *directionally complete* if it specifies in what direction one should go in order to reach the goal. Take for example the goal to substantially decrease the number of unemployed in Sweden. That goal is directionally complete because it suggests in what direction progress towards the goal is to be made. If employment has decreased, then the goal has been approached or achieved, otherwise not. A goal is *temporally complete* if it specifies the timeframe within which it should be attained.

A goal is *end-state evaluable*, moreover, if it is possible to know whether it has been achieved. The goal to reduce a pollutant in the atmosphere to a certain level that is far below what can be measured would fail to satisfy the criterion of end-state evaluability. A goal is *progressively evaluable* if it can be determined how far we are from satisfying it. This property of goals is crucial in determining whether a certain course of action should be maintained, changed or given up. It has also been argued that such feedback enhances the agent’s motivation so that she will make an intensified effort to act in ways that further the goal.

For an illustration, suppose my goal is to reach Geneva by the end of the day. In order for that goal to be rational, I must be able to determine whether or not this is the city I am actually in by the end of the day. However, in many situations it is not enough to be able to determine whether or not the goal state has been fully achieved. In the example, I must also be able to tell whether I am travelling in the right direction, and how far I have left to go. In particular, if

a goal is distant, or difficult fully to achieve we need to be able to judge the degree of success in approaching the goal. In other words, degrees of *partial attainment* must be distinguishable.

The second aspect of goal rationality concerns what the agent *can do*. It is reflected by the requirement that a goal should be *attainable*, or at least *approachable* (i.e. attainable at least to some degree). The goal to become a wizard (in the sense of a person with true magical powers) would not be classified as attainable or even approachable. There are at least three dimensions of approachability: *closeness*, *certainty* and *cost*. The dimension of closeness is the most obvious one. It concerns how close to the goal it is possible to come. The goal to achieve a perfectly just society is probably not fully achievable, and would therefore qualify as utopian, but it can be approached by acting in ways that increase social justice.

The third aspect of goal rationality is the volitional one. It concerns what we *want to do*. Goals, in order to be rational, should be motivating. Setting ourselves the goal should motivate us to act in a way which furthers the realization of the goal state. The motivation that a goal may give rise to in the agent can be characterized according to degree of intensity or durability. Studies indicate that goals are more action-generating when they are explicit and specific, and that such goals are more likely than do-your-best goals to intensify effort. There is also evidence suggesting that specific and challenging goals lead people to work longer at a task. We have already mentioned a connection between evaluation and motivation: when people can check how they stand in relation to a goal, their motivation to carry out the task often increases.

An insight into the nature of goal-setting emerging from SMART+ is that the criteria of rational goal-setting may conflict in the sense that the satisfaction of one criterion to a high degree may lead to a failure to satisfy substantially some other criterion. The probably most common type of such conflicts are occasioned by the fact that some of the properties that make a goal action-guiding may at the same time make it less capable to motivate action. Consider, for example, the following two goals (Edvardsson and Hansson, 2005):

- (1) The team shall win 12 out of 20 games with a least a two goal advantage, 3 out of 20 games with at least a one goal advantage, and never lose a game with more than one goal.
- (2) The team shall beat all opponents hands down.

Here, the second goal, though less action-guiding than the first, is plausibly more achievement inducing, and therefore more rational, because of its greater action-motivating capacity.

In general, visionary and utopian goals are more likely to motivate action than less visionary goals, which on the other hand may be more action-guiding. The task of goal-setting therefore may very well involve a trade-off between goals that are action-motivating and goals that are action-guiding. This may lead to the formulation of one single goal reflecting this compromise. However, it is often a better idea to adopt not a single goal but a whole system of goals at different levels. As Edvardsson and Hansson point out: “One way of balancing the criteria so as to optimize goal realization is to adopt goal systems in which goals are set on different levels in order to supplement each other. In this way visionary and highly motivating goals can be operationalized through more precise and evaluable subgoals, or interim targets.” (2005, p. 359)

While the SMART+ theory was developed for the purposes of studying managerial goal-setting rationality, it has been noted to be “well applicable to engineering practice” (Hughes, 2009, p. 393). There is one exception to this rule, though: the motivity of a goal, the degree to which it motivates those involved to work towards it, is not much discussed in the design literature, being more naturally regarded an issue for management.³ As for approachability, engineers pay special attention to this aspect via feasibility studies which is a standard tool for securing approachability. Precision is vital in engineering, not least in the problem definition phase. The design process starts with the identification of a need, whether it arises from a client’s needs or in some other way, giving rise to the broad goal of the engineering project. However, it is necessary that this goal be clarified before any serious design work can be done. Thus, the problem definition stage is often conceived of as an essential part of the design process (e.g. Dym and Little, 2004, Dieter, 1983). This clarificatory process involves considering design objectives, user requirements, constraints and functions. Moreover, engineers continuously verify that the design process meets the defined goals and that the goals are reasonable, giving rise to a feedback loop that may lead to further design refinements. Hughes concludes that “like precision, evaluability is a valued feature in engineering” (ibid.).

Another aspect of goal rationality concerns the coherence of ultimate goals. We would like to have, in management as well as in technology, ultimate goals that are mutually consistent and

³ The account of goal-setting in technology that follows is based on Hughes (2009), p. 393.

perhaps even mutually supportive, in the sense that the fulfillment of one goal facilitates the fulfillment of another (Rosencrantz, 2008). If two goals are mutually inconsistent, that means that both goals cannot be satisfied at once. Thus, a mousetrap designer would be irrational to insist that her product should not harm or restrict the freedom of the mouse in any way. But there may also be other more subtle forms of goal conflicts. Thus the satisfaction of one goal may make the satisfaction of another goal less likely. For instance, goal conflicts are common in environmental politics (Edvardsson, 2004). There are two kinds of goal conflicts within this context: internal and external. An example of an internal goal conflict is when a prohibition on using pesticides results in an increased mechanical use of land, which in turn increases the discharge of carbon dioxide. An external goal conflict arises when the expansion of the infrastructure in terms of roads and railroads threatened the preservation of sensitive biotopes. Softer forms of goal conflicts commonly arise in technology as well (Hughes, 2009). To take a frequently occurring case, success in reducing cost will often enough have an adverse effect on performance, and vice versa.

Few authors would go as far as disallowing goal conflicts altogether. The more common view is to regard them as a fact of life. Thus, Edvardsson concludes that “goal conflicts seem to be unavoidable in any multigoal system, unless the goals are set very low” (Edvardsson, 2004, p. 178), adding that “[t]he ideal of a perfect state of affairs in which all goals have been achieved is most likely utopian” (ibid.). It is not clear, in many such cases, that it is worthwhile to invest resources in formulating goals so that there cannot be a conflict. It may be a better strategy to allow for certain conflict potential and make the necessary trade-offs if a conflict should materialize in practice. Still, as Edvardsson points out, “it makes sense for any multigoal system to provide for some mechanism whereby goal conflicts may be solved” (ibid.). Contingency planning may be called for if conflicts are not only possible but can be expected to occur, although one should probably take into account that planning itself consumes resources. Hughes concludes, similarly, that goal conflicts should generally be allowed, noting that “[t]he [weak] consistency criterion is best reserved for those situations in which the goals are stated precisely (the cost should be no more than \$ x and the vehicle should accelerate at a rate of no less than y)” (Hughes, 2009, pp. 392-393).

3. Goal rationality in science

On first sight, goal rationality in science seems attractively simple in comparison to goal rationality in technology since the goal of scientific inquiry is simply to find the truth. On

closer scrutiny, however, considerable complexity emerges. For one, the goal to find the truth does not by itself suggest any very definite course of action; it does not specify in what direction one should go in order to reach the goal, except possibly that one should use a method that is reliable – one that is likely to lead to true beliefs. Still, the goal itself does not indicate what those methods are. Not only directional completeness but also the other aspects of goal rationality identified in the SMART+ model make good sense as principles governing goal rationality in the scientific domain. For another example, it would clearly be desirable in science to have a goal that is end-state evaluable in the sense that it is possible to know whether it has been achieved. Once more, the goal of truth is not an obvious candidate. Similarly, we would like scientific goals to be temporally complete, progressively evaluable, attainable, and we would be happy to have goals that exert the proper motivational force on the inquirer. Finally, there seems to be no reason to think that science is devoid of goal conflicts. For instance, the goal of truth could be satisfied by simply adopting a trivial theory, one which is logically true. To avoid this, we need the further goal of informativity. But as many epistemologists have observed (e.g. Levi, 1967), if we decide to adopt both goals as ultimate ends this is likely to lead to a goal conflict since a more informative theory is often less likely to be true. A theory that is very specific regarding the causes of a particular kind of cancer may thereby be less likely to be true than a less committed theory.⁴

To add substance to these remarks concerning the prima facie structural similarity between goal rationality in science and technology I will apply the SMART+ framework to four (related) epistemological debates concerning the proper goal of (scientific) inquiry, starting with two pragmatists: Peirce and Rorty. My main aim is to indicate, without focusing excessively on interpretational details, how the SMART+ theory could inform and clarify some otherwise puzzling epistemological disputes.⁵

4. Peirce on belief as the goal of inquiry

In a famous essay, Peirce argues that, contrary to the received view, the goal of inquiry is not truth, or true belief, but merely belief or “opinion” (Peirce, 1955, pp. 10-11):

⁴ For a related issue and some complications, see Bovens and Olsson (2002).

⁵ For the purposes of simplicity and definiteness, I will in the following take “truth” in its objectivist or realist sense as referring to correspondence with an external reality, although I conjecture that much of the reasoning that follows would survive a weakening to “empirical adequacy”, or the like.

[T]he sole object of inquiry is the settlement of opinion. We may fancy that this is not enough for us, and that we seek, not merely an opinion, but a true opinion. But put this fancy to the test and it proves groundless; for as soon as a firm belief is reached we are entirely satisfied, whether the belief be true or false. And it is clear that nothing out of the sphere of our knowledge can be our object, for nothing which does not affect the mind can be the motive for mental effort. The most that can be maintained is, that we seek for a belief that we shall *think* to be true. But we think each one of our beliefs to be true, and, indeed, it is mere tautology to say so.

We recall that, for Peirce, belief or opinion is, by definition, that upon which an inquirer is prepared to act. Hence, Peirce is proposing to reduce the goal of scientific inquiry to the goal of attaining that upon which we are prepared to act. Peirce's enterprise could be interpreted as presenting the TU-intuition with a direct challenge, especially if the kind of action that is referred to is practical action, a reading which is not alien to Peirce's pragmatism.

In the latter part of the quote, Peirce seems to be maintaining that the true state of things does not affect the mind and therefore cannot be the motive of mental effort. But the claim that the facts of the matter do not affect the mind is a counterintuitive one. When I look out the window, I come to believe that there is a tree just 10 meters away. Normally, this belief is caused by the tree, or the fact that there is a tree, which is thus affecting my mind.⁶

On another interpretation, Peirce is thinking of objective truth as essentially "mind-independent". If so, one could be led to think that it follows trivially that objective truth cannot affect the mind, for nothing that is mind-independent can if that is what "mind-independent" means. But this is an irrelevant sense of mind-independence. In a less trivial sense, something is mind-independent and objective if it does not depend entirely on our will. Truth is mind-independent in the latter sense but not in the former. What is true – for example that there is a tree outside the window – does not depend entirely on our will but it is still something that can affect us in various ways, and typically does so through our observations.

Peirce is right, though, in stating that once we believe something, e.g. that there is a tree out there, we cannot, pending further inquiry, distinguish the state we are in from a state of *true* belief. If S believes that p, or believes truly that p, she cannot tell whether she has attained the

⁶ It could be objected that Peirce is here using "truth" in a technical sense, signifying what is collectively accepted by all researchers once scientific inquiry has come to an end. Truth in that sense presumably does not exert any direct influence on a particular mind now. Still, this is an implausible interpretation of Peirce in the present context, as there is no concrete sign that truth should be given any special technical meaning.

first goal or the second. She will, from the position of the goal end state, judge that she believes that p just in case she will judge that she believes truly that p . Peirce can be understood as maintaining that this fact alone makes it more rational, or appropriate, to view the goal of inquiry in terms of fixing belief rather than in terms of fixing *true* belief. Is that correct?

Let us look at the matter from a more abstract perspective. We will say that two goals G_1 and G_2 are *end-state evaluation equivalent* for a subject S if, upon attaining one of G_1 or G_2 , S cannot tell whether she attained G_1 or G_2 . Peirce, in the argument under scrutiny, is relying on the following principle:

(Peirce's Principle) If (i) G_1 and G_2 are end-state evaluation equivalent for a subject S , and (ii) G_1 is logically stronger than G_2 , then G_2 is more rational than G_1 for S .

Is this principle valid as a general principle of goal rationality? I will argue that it is not. Suppose that P is a pollutant that is dangerous to humans and that M is a device which indicates whether or not the amount of P in the air exceeds the limits that have been set by an international body. Moreover, there is no other device that can be used for this purpose. However, M is not fully reliable and it sometimes misfires. Let G_1 be the goal of using the device M *and* successfully determining whether the air is free of P -pollution; and let G_2 be the goal of using the device M . G_1 and G_2 are end-state evaluation equivalent for the measuring person S : upon attaining G_1 or G_2 she cannot distinguish one from the other. Moreover, G_1 is logically stronger than G_2 . It would follow from Peirce's Principle that G_2 is more rational than G_1 . But this conclusion can be questioned. It is true that G_2 is more easily attained than G_1 . But G_1 is surely more inspiring than G_2 ; it is, to use Peirce's own expression, a stronger "motive for mental effort". It cannot, therefore, be concluded that G_2 is more rational, or achievement-inducing, than G_1 . Hence, the principle presupposed by Peirce is plausibly not generally valid. This observation is sufficient to undermine Peirce's argument that the goal of belief is more rational, or appropriate, than the goal of true belief.

Indeed, the goal of true belief, or the goal of truth for short, does sound more inspirational than the goal of settling belief. Many people, not least those equipped with a scientific mind, will go to almost any length to find the truth of the matter, sometimes even in practically insignificant affairs. Disregarding the special case of religious faith, comparatively few would be willing to incur similar personal and other costs for the sole gain of settling a corresponding opinion.

Apart from the general invalidity of Peirce's Principle, there may be other differences between the goal of belief and that of true belief that are worth attending to. One such factor is a difference in precision. We recall that a goal is said to be directionally complete if it specifies in what direction one should go in order to reach the goal. We have noted that the goal of truth does not do terribly well on this score. But it might still do better than the goal of belief. For the goal of true belief suggests, albeit imperfectly, that the belief be fixed, not by any old method, but by one that is likely to establish the truth of the matter. This would suggest to the inquisitive mind such things as evidence-gathering, hypothesis-testing, the use of scientific instruments, and so on. The goal of belief does not suggest as vividly any particular course of action. It is compatible with using a wider range of methods, including methods that are not truth-oriented but focus, say, on the systematic disregard of contravening evidence.

Finally, there is a difference between the two goals on the ability dimension, concerning what we can do to approach the respective goals. This is related to the presumed difference in directional completeness. The goal of belief can be approached and evaluated along one dimension only: degree of belief. The stronger our belief is, the closer we are to achieving the goal of (full) belief. The goal of truth, by contrast, can in addition be approached, at least in principle, along the dimension of truth-likeness: the closer we are to the truth, the closer we are to achieving the goal of true belief *ceteris paribus*.

5. Rorty on justification as the goal of inquiry

My second application of SMART+ concerns an argument presented by Richard Rorty in a paper from 1995, drawing partly on earlier work (e.g. Rorty, 1986), to the conclusion that truth is not legitimately viewed as the goal of inquiry. This is a conclusion also drawn by Peirce, as we saw, but where Peirce thought that the goal of truth should be replaced by the goal of belief, Rorty proposes that the proper replacement is rather *justified* belief.

The starting point of Rorty's 1995 is the following declaration (p. 281):

Pragmatists think that if something makes no difference to practice, it should make no difference to philosophy. This conviction makes them suspicious of the philosopher's emphasis on the difference between justification and truth. For that difference makes no difference to my decisions about what to do. If I have concrete, specific doubts about whether one of my beliefs is true, I can resolve those doubts only by asking whether it is

adequately justified – by finding and assessing additional reasons pro and con. I cannot bypass justification and confine my attention to truth: assessment of truth and assessment of justification are, when the question is about what I should believe now (rather than about why I, or someone else, acted as we did) the same activity.

He adds, a few pages later on (p. 286):

The need to justify our beliefs and desires to ourselves and our fellow agents subjects us to norms, and obedience to these norms produces a behavioral pattern that we must detect in others before confidently attributing beliefs to them. But there seems no occasion to look for obedience to an additional norm – the commandment to seek the truth. For ... obedience to that norm will produce no behavior not produced by the need to offer justification.

In arguing that the goal of scientific inquiry is not truth but being in a position to justify one's belief, Rorty is, in effect, challenging the TU-intuition, especially as he views justification as essentially unrelated to truth, which in the end is a notion he favors dropping altogether (p. 299). One of the conclusions of his essay is that, on the Dewey-inspired theory which he advocates, "the difference between the carpenter and the scientist is simply the difference between a workman who justifies his action mainly by reference to the movements of matter and one who justifies his mainly by reference to the behavior of his colleagues" (ibid.).

My ambition here is not to add to the voluminous literature on the interpretation of Rorty's pragmatism. Instead, I would like to distill one argument that I believe can be found in his essay, suitably reconstructed. Rorty, as quoted above, is contrasting two goals: the goal of attaining a true belief and the goal of attaining a justified belief. On the reading I would like to highlight, he is offering an argument that is similar to Peirce's argument for the propriety of the goal of belief, but for a slightly different conclusion. Rorty is pointing out that the goal of attaining a true belief and the goal of attaining a (sufficiently) justified belief are end-state evaluation equivalent from the point of view of the inquirer: once the inquirer has attained either of these goals, she cannot tell which one she attained. This much seems true. Yet Peirce's Principle is not directly applicable as it demands that, among the goals under consideration, one goal be logically stronger than the other. The two goals of true belief and

justified belief are not at all logically related, at least not as justification is standardly conceived.⁷

Still, we note that the goal of justified belief is plausibly more directionally complete than the goal of true belief, and in the quote this is a feature that Rorty highlights. On a plausible reconstruction, the general principle underlying Rorty's reasoning, then, is this:

(Rorty's Principle) If (i) G_1 and G_2 are end-state evaluation equivalent for a subject S , and (ii) G_2 is more directionally complete than G_1 , then G_2 is more rational than G_1 for S .

But this principle shares the fate of Peirce's Principle of being plausibly generally invalid. Since the problem is similar in both cases, I shall not this time give an explicit counterexample. Suffice it to note that beside directional completeness, there are several other aspects of a goal that play a part in determining its relative rationality. One such aspect is, to repeat, the motivational one. This aspect is interesting in this context because it often offsets the directional aspect. Goals that are strongly motivational are in practice rarely directionally complete, and vice versa. Thus many are motivated by goals such as achieving "world peace" or "a completely just society" and yet these goals do not *per se* suggest any particular cause of action. Conversely, goals that give detailed advice for how to act tend to be less inspirational.

As we have already noted, the goal of truth, though directionally less complete than the goal of justification, may still be more rational in virtue of its inspirational qualities. Hence, *pace* Rorty we cannot conclude, from the presumed fact that the goal of true belief and the goal of justified belief are end-state evaluation equivalent and the latter more directionally complete than the former, that the latter is also the more rational choice.

Leaving Rorty's discussion aside, a natural view to adopt concerning the relation between the two goals of true belief and justified belief, from a SMART+ perspective, is that they could very well live side by side, supplementing each other: the goal of truth providing the visionary, motivating factor and the goal of justification playing the more action-guiding part. Drawing on the upshots of section 3, there are *prima facie* two ways of implementing this recommendation. One would be to adopt a system of goals wherein both goals figure, the goal of truth as a high-level goal and the goal of justification as lower-level goal, the latter operationalizing the former. The other way would be to compress the two goals into one goal, the goal, namely, to attain a justified true belief. The latter goal amounts, incidentally, to the

⁷ For more on this, see the section on Kaplan below.

goal of attaining *knowledge*, as that concept is traditionally conceived. Yet the claim that knowledge, in the traditional sense, is the proper goal of inquiry has been questioned by several epistemologists. I turn now to two such criticisms, due to Mark Kaplan and Crispin Sartwell, respectively.

6. Kaplan on the irrelevance of knowledge in inquiry

Fifty years ago, Edmund Gettier famously argued that the traditional analysis of knowledge is mistaken (Gettier, 1963). Gettier proposed two counterexamples intended to show that we may be justified in believing something which is true, without this belief qualifying as a case of knowledge in a pre-systematic sense. Gettier's paper generated an industry of attempts to solve the "Gettier problem", which is standardly interpreted as the problem of identifying additional clauses which, if they are added to the traditional account of knowledge, make Gettier's examples cases of non-knowledge. For an overview, the reader is referred to Shope (2002).

A rather different approach was taken by Mark Kaplan (1985). Rather than putting forward further clauses to supplement the traditional account knowledge, Kaplan challenged the importance of the Gettier problem as such. If it is not an important problem, then there is no pressing need to solve it. The aim of this section is to examine Kaplan's arguments more closely, and to do so specifically from the perspective of goal-setting rationality.

To fix ideas, one of Gettier's counterexamples takes the following form. Both Smith and Jones have applied for a certain job. Smith is justified in believing both that Jones will get the job and that Jones has ten coins in his pocket. Because of this, Smith also justifiably believes that the man who will get the job has ten coins in his pocket. As it turns out, unbeknownst to Smith, it is he, not Jones, who will get the job. And, as it happens, Smith, too, has ten coins in his pocket. Smith, then, has a justified true belief that the man who will get the job has ten coins in his pocket. But given the circumstances it seems that Smith lacks knowledge of this fact. If so, having a justified true belief is not sufficient for having knowledge.

As Kaplan notices, the solution to the problem may look obvious. The reason why Smith lacks knowledge of the proposition in question is surely the fact that his conclusion relies on a false premise, that, namely, Jones will get the job. Hence, the problem can be avoided by simply adding a further clause to the traditional account of knowledge ruling that, in order for a subject to know that *p*, the subject's reasons for *p* must not rely on a false premise.

Unfortunately, there are similar examples which do not involve reasoning from a false premise (Goldman, 1976). For our purposes, we may disregard this complication. We will follow Kaplan and assume, for the sake of the argument, that the no-false-premise solution properly handles the Gettier problem.

Now the Gettier problem is a problem for the analysis of knowledge as justified true belief only if we conceive of justification as being “fallible”. In other words, we need to suppose that a subject can be justified in believing that *p* even if *p* is, in fact, false. In Gettier’s example it was assumed that Smith justifiably believes the false proposition that Jones will get the job. But – and this is Kaplan’s first point – it can be questioned whether historically influential philosophers were fallibilists regarding justification. Descartes, for one, is usually taken rather to subscribe to infallibilism. For him, justification, which he took to involve the clear and distinct grasping of the truth of a proposition, cannot obtain unless the proposition that is thus justified is in fact true. From Descartes’ perspective, then, there could not be a Gettier problem. This is the first reason why Kaplan thinks that the Gettier problem is not as important as it is commonly taken to be; contrary to popular opinion, it does not challenge an account of knowledge that is properly called “traditional”.

We will leave this aspect of Kaplan’s argument aside and focus on the second element of his reasoning, which is more relevant here because it involves, at least implicitly, considerations of goal setting. Kaplan’s second reason for downplaying the Gettier problem stems from his particular account of what it takes for a philosophical problem to be significant. Surveying some historically important debates, such as the logical positivists’ concern with a verificationist theory of meaning, he concludes that what makes a problem important is the extent to which solving it “succeeded in advancing or clarifying the state of the art of inquiry” (p. 354). But it is not clear, he continues, how solving the Gettier problem would yield insights in this regard.

Kaplan invites us to consider, as a preliminary, the following issue for the so-called traditional account: Suppose that you as a responsible inquirer have considered all the available evidence, whence you conclude that *p*. Having done so, there appears to be no further point in asking whether you also *know* that *p*. Once you have satisfied yourself that you have a justified belief, you have thereby also satisfied yourself that your belief is true:

From where you sit, determining whether you believe *p* with justification and determining whether you know that *p* comes to the same thing. But then, far from being

integral to your pursuit of inquiry, distinguishing the propositions you know from those you don't know is, on the justified-true-belief analysis, a fifth wheel. (p. 355, notation adapted)

Kaplan proceeds to argue that the same issue arises for the post-Gettier account of knowledge:

In so far as you are satisfied that your belief in *p* is well founded, you will ipso facto be satisfied that you have not inferred *p* from a false premise – otherwise you would *not* think you had good reason for concluding that *p*. Just as on the justified-true-belief analysis, determining whether you believe *p* with good reason and determining whether you know that *p* come to the same thing. (p. 355, notation adapted)

Kaplan's conclusion is that "what you know doesn't matter" (p. 362).

Kaplan's argument can be reconstructed as one based on considerations of goal rationality. From this perspective, the first part – concerning the traditional concept of knowledge – is practically identical to Rorty's (later) argument to the conclusion that truth is not a goal of inquiry; only justification is. Kaplan is, in effect, appealing to what I have called Rorty's Principle: if (i) two goals G_1 and G_2 are end-state evaluation equivalent for a subject *S*, and (ii) G_2 is more directionally complete than G_1 , then G_2 is more rational than G_1 for *S*.⁸ We saw that this principle is not plausible from a goal theoretical perspective because it fails to take into account the motivational aspect of visionary goals such as the goal of truth. Since the second part, about Gettier, presupposes the first part, it is just as unconvincing.

One could add that while the goal of attaining a justified belief where the justification is not based on false premises may seem to coincide with the goal of attaining a justified belief *simpliciter* because, in Kaplan's words, "none of us needs to be taught that an argument with a true conclusion does not carry conviction if that conclusion rests upon a false premise" (p. 359), there are circumstances in which the former may have some action-guiding qualities that the latter lacks. For instance, if the inquirer is investigating an emotionally loaded issue, in which she has strong personal stakes, a reminder that her position is not stronger than her

⁸ Since Kaplan's paper appeared ten years before Rorty's, calling the principle in question "Kaplan's principle" would do more justice to the actual chronology. However, while Kaplan's argumentation is on the whole clearer than Rorty's, it must be said, to his credit, that Rorty is more explicit about the particular fact that he is addressing a problem of rational goal setting. Also, as I mentioned, Rorty's 1995 paper reflects ideas that he has expressed in earlier works, including works from the 1980's.

weakest premise may prevent her from wishfully adopting premises that support her favored conclusion.

7. Sartwell on knowledge as mere true belief

In his paper from 1992, Crispin Sartwell seeks to establish the, for a traditional epistemologist, surprising thesis that knowledge is mere true belief. His argument proceeds from the premise that knowledge is the overarching or ultimate goal of inquiry. What we seek in inquiry is, above all, knowledge. For why else, he asks, should knowledge occupy such a central place in epistemology?

Now suppose that knowledge, *pace* Kaplan, is justified true belief. There are two ways of conceiving the value of justification. Either it derives wholly from the value of true belief or it does not do so. Consider the first possibility. In that case, justification is merely a criterion of some part of the goal of inquiry, namely true belief, and cannot therefore itself be part of that goal, or so Sartwell believes. Consider instead the other possibility, i.e. that of justification being valuable independently of truth. In that case, Sartwell thinks, knowledge is an incoherent concept combining two independently valuable components which cannot always be realized simultaneously. But knowledge is not an incoherent concept. Hence, knowledge is mere true belief.

Sartwell is here assuming, in the first part of his argument, that an ultimate goal cannot contain parts which are valuable only as criteria of other parts of that goal. That is the reason why he thinks justification, if it is merely a criterion of truth, cannot be part of knowledge, if knowledge is conceived as involving truth and as being at the same time the ultimate goal of inquiry. What can be said of this part of the argument from the point of view of the general theory of goal-setting? It is indeed in the spirit of SMART+ to separate ultimate from instrumental goals in the description of a goal system, although as far as I can see, that framework does not strictly speaking disallow ultimate goals that have instrumental goals as parts.

It is rather in the second part of his argument that Sartwell may have failed more substantially. Consider his argument for thinking that if justification is valuable independently of its relation to truth, then knowledge, in the sense of justified true belief, is an incoherent and therefore useless notion. To be specific, he levels the following objection against William Lycan's proposal that justification has certain independent virtues related to explanatory elegance:

[K]nowledge turns out to be (at least) true belief that is generated by adaptive explanatory techniques. But this seems odd: now that we recognize two primitive epistemic values, they may well conflict. For example, is it good to believe, in some circumstances, highly explanatory falsehoods? The account surely leaves some such cases strictly undecidable, since it describes both elegance and truth as intrinsic values. But is this plausible? Surely, we might want to say, though it can be useful to believe all sorts of falsehoods, it is always epistemically good to believe the truth. It may be useful, for example, for me to have cognitive technique that causes me to believe that I have all sorts of positive qualities to an extremely high degree ... But it is not a good thing epistemically to believe such things if they are false.

And he proceeds: “But we cannot coherently demand that we follow both of these as ultimate aims, because they may and in fact will conflict. Then we are left with an internally incoherent concept of knowledge.”

The critical general issue here is whether the fact that two goals may sometimes conflict shows that they cannot both be ultimate ends. Sartwell thinks that this is indeed the case. However, as we saw earlier there are numerous examples of goal systems in management or technical contexts listing as ultimate aims goals that can, or even can be expected to, conflict. To take an example, the overall goal of the Swedish transport system is in fact a combination of two goals: economic efficiency and long-term sustainability. It is quite easy to imagine circumstances in which both goals cannot be attained at once. We need only imagine a case in which the economically most efficient system, because of the natural resources it consumes, is one which can only operate for a relatively short period of time. We recall that environmental politics is another area in which potential goal conflicts abound.

What follows for the purposes of Sartwell’s argument? First of all, there is nothing wrong per se in having more than one ultimate goal that may conflict, such as the two goals of attaining the truth and attaining beliefs of high explanatory value. But since goal conflicts can be expected to occur, it would make sense to add a meta-rule specifying what to do when there is a tension. Such a rule could for instance stipulate that the goal of truth is to take priority in such cases over the goal of explanatory value.

The bottom line is that a proponent of the justified true belief account of knowledge can still hold that knowledge is the goal of inquiry, not least if she is willing to tie justification not to truth but to some other value independent of truth, or alternatively conceive of justification as

itself intrinsically valuable. I am not recommending this move but only pointing out that Sartwell is unsuccessful in ruling it out, and – importantly – that this is so for reasons that have to do with general principles of goal rationality. In any case, taking this path may well be combined with an effort to devise a precautionary strategy for handling expected goal conflicts.

8. Conclusion

The question raised in this paper was to what extent goal setting rationality in technology is different from such rationality in science. My thesis, or conjecture, was that there is no substantial difference. The principles underlying the SMART+ theory are just as plausible for scientific goal setting as they are for technological goal setting. Moreover, the systematic advantages in viewing SMART+ as a unified theory of rational goal setting, covering all such activity, are substantial, not least for the purposes of epistemology. I tried to mount support for the latter claim by reconstructing four related epistemological debates concerning the proper goal of inquiry in the light of this unification thesis. My thesis was that all four debates depend on principles of rational goal setting that are not, or do not appear to be, generally valid. The most common oversight, from this perspective, is the failure in epistemology to take into account the motivational aspect of visionary goals, most prominently, the goal to attain objective truth. Curiously, this failure seems deeply rooted in pragmatist writings – we saw it in the writings of both Peirce and Rorty – without its apparent incompatibility with other features of pragmatism being clearly brought to the fore. I am thinking obviously of the pragmatist claim that what matters in philosophy is what makes a practical difference, from which it is concluded – to make a long story short – that truth cannot be a goal of inquiry. But the fact of the matter is that the goal of truth should rather be cherished by pragmatists as a goal which, due to its tendency to move inquirers to increase their mental effort, is as practice-affecting as one could have wished.

What also comes out, potentially, of this study is that the epistemological focus on identifying *the* goal of inquiry appears unmotivated and even somewhat obsessive. From the current perspective, there are reasons to focus less on finding a unique goal of inquiry and more on finding a plausible *system* of such goals. Finally, there is a tendency in the epistemological literature to think that potential goal conflicts should be avoided at all costs, as evidenced by Sartwell's article. But on the more general picture of goal rationality advocated here, potential goal conflicts are a fact of life, and the only conclusion that follows from the fact that goal

conflicts can be expected to occur is the practical one that the designer of the goal system may want to invest some resources in planning ahead for the various contingencies that may materialize in the future.

References

Bovens L, Olsson EJ (2002), Believing more, risking less: on coherence, truth and non-trivial extensions. *Erkenntnis* 57:137-150

Dieter, GE (1983) *Engineering design: a materials and processing approach*. McGraw-Hill, New York

Dym CL, Little P (2004) *Engineering design: a project-based introduction*. John Wildy and Sons, New York

Edvardsson K (2004) Using goals in environmental management: the Swedish system of environmental objectives. *Environmental Management* 34(2):170-180

Edvardsson K, Hansson SO (2005) When is a goal rational? *Social Choice and Welfare* 24:343-361

Gettier E (1963) Is justified true belief knowledge? *Analysis* 23:121-123

Goldman AI (1976) Discrimination and perceptual knowledge. *The Journal of Philosophy* 73:771-791

Houkes W (2009) The nature of technological knowledge. In: in Meijers A (ed) *Philosophy of technology and engineering sciences*. *Handbook of the philosophy of science*, vol 9. Elsevier, Amsterdam, pp 309-350

Hughes J (2009) Practical reasoning and engineering", In: in Meijers A (ed) *Philosophy of technology and engineering sciences*. *Handbook of the philosophy of science*, vol 9. Elsevier, Amsterdam, pp 375-402

Kaplan M (1985) It's not what you know that counts. *The Journal of Philosophy* 82(7):350-363

Levi I (1967) *Gambling with truth: an essay on induction and the aims of science*. Routledge and Kegan Paul, London

- Mali P (1972) *Managing by objectives: an operating guide to faster and more profitable results*. John Wiley and Sons, New York
- Peirce CS (1955) The fixation of belief. In: Buchler J (ed) *Philosophical writings of Peirce*. Dover Publications, New York, pp 5-22 (First published in *Popular Science Monthly*, 1877)
- Rorty R (1986) Pragmatism, Davidson and truth. In: LePore E (ed) *Truth and interpretation: perspective on the philosophy of Donald Davidson*. Basil Blackwell, Oxford, pp 333-368
- Rorty R (1995) Is truth a goal of inquiry? Davidson vs. Wright. *The Philosophical Quarterly* 45(180):281-300
- Rosencrantz HK (2008) Properties of goal systems: consistency, conflict, and coherence. *Studia Logica* 89:37-58
- Rosencrantz HK, Edvardsson K, Hansson SO (2007) Vision zero – is it irrational? *Transportation Research Part A: Policy and Practice* 41(6):559-567
- Sartwell C (1992) Why knowledge is merely true belief. *The Journal of Philosophy* 89(4):167-180
- Shope RK (2002) Conditions and analyses of knowing. Moser PK (ed) *The Oxford handbook of epistemology*. Oxford University Press, Oxford, pp 25-70
- Simon HA (1956) Rational choice and the structure of the environment. *Psychological Review* 63(2):129-138