3
The Weight of Argument

Isaac Levi

3.1 The somewhat novel question

In Chapter 6 of *A Treatise on Probability* (1973 [1921], hereafter *TP*), John Maynard Keynes raised a question he described as "somewhat novel." This concerned the problem he labeled "The weight of arguments."

Keynes claimed that the magnitude of the probability of an argument depends upon "a balance between what may be termed the favourable and the unfavourable evidence" (*TP*, p. 77). In contrast to this, Keynes introduced a notion of weighing the amount of available evidence.

3.2 Peirce on balancing reasons

Charles Peirce had exploited the metaphors of balancing and weighing evidence in 1878 in his "Probability of Induction." Unlike Keynes, however, he did not contrast "balance" and "weight" but used them more or less interchangeably.

Peirce devoted considerable space in that essay to examining the "conceptualist" view of probability, according to which probability "is the degree of belief that ought to attach to a proposition" (Peirce, 1878, p. 291).

According to conceptualists, degree of belief can be measured by a "thermometer" whose reading is proportional to the "weight of evidence" (Peirce, 1878, p. 294). To explain this, Peirce focused on cases where one is assessing evidence for and against a hypothesis H and where each datum is probabilistically independent of every other relative to information prior to data collection. He derived the probability that two items testify in favor of H conditional on both items agreeing in testifying either for or against H. He then defined the "chance of H" as what

S. M. D. Brandolini et al. (eds.), *Fundamental Uncertainty*

is now called the "odds for H given E" or ratio of P(H/E)/P(~H/E) where E describes the data. He took the conceptualist practice of invoking insufficient reason to be claiming that the prior odds P(H)/P(~H) = 1 and with that took the "final odds" (that is, the odds for H given E) to be equal to the "likelihood ratio" P(E/H)/P(E/~H). He then argued that the logarithm of this ratio is a suitable "thermometer" for measuring degrees of belief. This logarithm is equal to log P(E/H) − log P(E/~H). E is a reason for H if this value is positive. E is a reason against H if the value is negative. Peirce also added:

> But there is another consideration that must, if admitted, fix us to this choice for our thermometer. It is that our belief ought to be proportional to the weight of evidence, in this sense, that two arguments which are entirely independent, neither weakening nor strengthening each other, ought, when they concur, to produce a belief equal to the sum of the intensities of belief which either would produce separately (Peirce, 1878, p. 294).

In spite of Irving John Good's allegations to the contrary (Good, 1981), Peirce's characterization of the "independence" of two arguments concurring in their support for or against the hypothesis H is perfectly correct. If E = E_1&E_2 and independence here means that P(E/H) = P(E_1/H) P(E_2/H) and P(E/~H) = P(E_1/~H)P(E_2/~H), the logarithms of the products become sums. The likelihood ratio becomes a sum of two differences logP(E_1/H) − logP(E_1/~H) and logP(E_2/H) − logP(E_2/~H). If the two differences concur—that is, show the same sign—Peirce's contention is that the Fechnerian intensities of belief ought to be the sum. The "weight" of the support has increased. If the two differences are of different signs so that one bit of evidence supports H and the other undermines it, a "balancing of reasons" assesses the overall support (Peirce, 1878, p. 294). Peirce does not say so in so many words but we may call the result an assessment of the "net weight" of the reasons or evidence.

Peirce explicitly associated the balancing of reasons with the conceptualist view of probability that he *opposed*. He took the balancing reasons procedure to be a way of presenting the best case for conceptualism. Peirce pointed out that the approach presupposes a dubious use of the principle of insufficient reason to identify "prior" probabilities for use with Bayes's theorem to obtain posterior probabilities and final odds.

Peirce insisted that the best case is not good enough. Peirce's writings are strewn with diverse arguments attacking the conceptualist view.

Some of the attacks are against the principle of insufficient reason. Some of them are attacks on conceptualism even when it dispenses with insufficient reason and takes a more radically subjectivist turn. Peirce wrote:

> But probability, to have any value at all, must express a fact. It is, therefore, a thing to be inferred from evidence.
>
> (Peirce, 1878, p. 295)

There is some evidence to think that neither Peirce nor John Venn (whom Peirce clearly admired) always interpreted judgements of probability to be judgements of fact—that is, relative frequency or physical probability. Peirce was plainly concerned with issues where a decision maker is faced with a momentous decision in the here and now without the prospect of referring the outcome to a long-run series of outcomes. Judgements of probability used in the evaluation of prospects facing the decision maker are not beliefs about relative frequencies or physical probabilities.

According to Peirce, the probability judgements the inquirer is prepared to make in a given context of choice should be *derived* from the information the inquirer has concerning long-run relative frequencies or physical probabilities in situations of the kind the decision maker takes himself or herself to be addressing. Judgements of numerically determinate belief probability are worthless as a warrant for assessing risks (they have no value at all as Peirce puts his point) unless they "express" a fact. As I understand Peirce, the judgements of numerically determinate belief probability are derived in accordance with principles of "probabilistic syllogism," "statistical syllogism," or, as later authors might put it, "direct inference" from information about objective statistical probabilities of outcomes of an experiment of some kind. That an experiment of type S is also of type T is legitimately ignored if the statistical probabilities of outcomes of trials satisfying both S and T are *known* (or fully believed) to be equal to the statistical probabilities of those outcomes on trials of kind S. (The extra information that the trial is of kind T is *known* to be "statistically irrelevant.") If the inquirer knows the extra information to be statistically relevant or *does not* know it to be irrelevant, it may not be ignored even if this means that no determinate belief probability may be assigned to a hypothesis about the outcome of experiment.

Appealing to insufficient reason to assign probabilities is a way of deriving belief probability judgements that fail to "express" a fact.

For Peirce such derivation would be unacceptable. His opposition to conceptualism was opposition not so much to belief probabilities per se but to the appeal to considerations such as insufficient reason to ground the assessment of belief probabilities.

Peirce offered an illustration of his point of view and in that setting mounted an argument against the "proceeding of balancing reasons."

A bean is taken from a large bag of beans and hidden under a thimble. A probability judgement is formed of the color of the bean by observing the colors of beans sampled from the bag at random with replacement. Peirce considered three cases: (i) two beans are sampled with replacement where one is black and one is white; (ii) ten beans are sampled where four, five, or six are white; and (iii) 1000 are sampled and approximately 500 are black.

The conceptualist (known to us as Laplacian or Bayesian) invokes insufficient reason to assign equal prior probability to each of the $n+1$ hypotheses as to the number r of white beans in the bag of n beans. Bayes's theorem yields a "posterior distribution" over the $n+1$ hypotheses. One can then obtain an estimate of the average number r of white beans in the bag of n that is equal to the probability on the data that the bean under the thimble is white. Peirce found this derivation acceptable in those situations where the inquirer could derive the prior probability distribution from knowledge of statistical probabilities. To achieve this in every case, the inquirer would have to assume absurdly that worlds are as plentiful as blackberries.

The alternative favored by Peirce is to consider a rule of inference that specifies the estimate to make of the relative frequency of whites in the bag for each possible outcome of sampling. Indeed, if the rule specifies that the estimate of the relative frequency of whites in the bag falls in an interval within k standard deviations from the observed relative frequency of whites in the sample, it can be known prior to finding out the result of sampling what the statistical probability of obtaining a correct estimate will be. So, in making a judgement about the color of the ball under the thimble, we can no longer take the excess of favorable over unfavorable data as a thermometer measuring degree of belief. Indeed, no single number will do.

In short, to express the proper state of our belief, not *one* number but *two* are requisite, the first depending on the inferred probability, the second on the amount of knowledge on which that probability is based (see Peirce, 1878, p. 295).

In a footnote to this passage Peirce made it clear that "amount of knowledge" is tied to the probable error of the estimate of the "inferred probability." Indeed, Peirce suggested that an infinite series of numbers may be needed. We might need the probable error of the probable error, and so on.

Thus, the "amount of knowledge" or what I am calling the "gross weight of evidence," according to Peirce, cannot be accounted for on the conceptualist view but can be on the view that insists that belief probabilities be derivable via direct inference from statistical probability.

3.3 Keynes as a conceptualist

Peirce would have classified Keynes as a conceptualist. For Keynes "the terms *certain* and *probable* describe the various degrees of rational belief about a proposition which different amounts of knowledge authorize us to entertain" (*TP*, p. 3.) To call the degrees of belief "rational" is to indicate that the "degree of belief" is that degree to which it is rational for *X* to entertain an hypothesis given the information or evidence available to *X*.

X is supposed to have a body of "direct" knowledge of logical relations between propositions that includes deductive entailments and judgements concerning the probability relations between data and hypothesis as well general principles characterizing the consistency of probability judgements with each other.

If we strip away the dubious elements of Russellian epistemology that surface in the first two chapters of Keynes's *Treatise on Probability*, the principles of probability logic as Keynes conceived it constitute constraints on the rational coherence of belief. Principles of deductive closure and consistency cover judgements of certainty. Principles of probability logic are constraints of rational coherence imposed on judgements of degree of belief.

Thus, Keynes suggests that if *X* has a set of exclusive and exhaustive hypotheses given *X*'s background information (or certainties, evidence, or knowledge), and if "[t]here [is] no *relevant* evidence relating to one alternative, unless there is *corresponding* evidence relating to the other," one should assign equal probabilities to each alternative (*TP*, p. 60). Keynes suggested that the judgements of relevance and irrelevance of the evidence required to apply the principle should be based on direct judgement. Such judgements seem to be judgements of a relation of probability between evidence and hypotheses of the sort that Keynes compared to the relation of deductive entailment. Frank Plumpton Ramsey (1990a, pp. 57–9) justly worried about whether one could ground such judgments directly. Nonetheless, Keynes's principle of insufficient reason or "Indifference," as Keynes called it, is a constraint on the coherence or consistency of such judgements of relevance and judgements of equality and inequality in probability judgement. As such it could be considered a "logical principle" in a sense akin to that according to which the requirement that judgements of probability obey the calculus of probability was taken by Ramsey to be a logical principle belonging to a "logic of consistency" for probability judgement.

Keynes's second principle requires the probability of *ab* given *h* to be less than the probability of *a* given *h* unless the probability of *b* given *ah* equals 1. Keynes offered a third principle stating that the probability of *a* given *h* is comparable with the probability of *a* given hh_1 as long as h_1 contains no independent parts relevant to *a*.

The important point to notice here is that Keynes thought of probability logic as imposing constraints on coherent probability judgement that did not always or, indeed, typically require rational agents to adopt a unique probability distribution over some domain given the evidence. Many systems of quantitative probability judgement might satisfy the constraints given the inquirer's "evidence" or body of certainties. Sometimes numerically determinate probability judgements are mandated. For example, when the principle of insufficient reason is applicable, a rational agent is constrained to adopt degrees of belief on his or her evidence that is representable quantitatively. But probability logic cannot constrain quantitative probability judgement uniquely. According to Keynes, as I understand him, the inquirer is then obliged as a rational agent to adopt a state of probability judgement representable by the set of all the probability functions permissible according to the constraints or, alternatively, by the judgements of comparative probability judgement that are implied by all such logically permissible probability judgements. The principles of probability logic need not mandate numerical degrees of belief that *a* and that *ab* on the evidence *h* but only that *X* is required to be no less certain that *a* than that *ab*. According to Keynes, probabilities of hypotheses on given information could even be non-comparable.[1]

Here then is one point of agreement between Peirce and Keynes: belief probability judgements can be and often are indeterminate. If we ask, however, how the indeterminacy arises in rational probability judgement, Peirce would respond by saying that whenever there is not sufficiently precise information about statistical or physical probabilities on which to base a derivation of belief probabilities via direct inference (and the calculus of probabilities), probability judgement should be indeterminate. Keynes, as I understand him, insisted that his principles, in particular Insufficient Reason or Indifference, could often warrant assigning determinate probabilities on evidence even though indeterminacy can prevail when the requisite conditions are not met.

3.4 Keynes on frequentism

In spite of their differences, Keynes agreed with Peirce that the belief probability judgements one can reach by balancing reasons cannot tell the entire story of how evidence is assessed:

> The magnitude of the probability of an argument . . . depends upon a balance between what may be termed the favourable and the unfavourable evidence; a new piece of evidence which leaves the balance unchanged, also leaves the probability of the argument unchanged. But it seems that there may be another respect in which some kind of quantitative comparison between arguments is possible. This comparison turns upon a balance, not between the favourable and the unfavourable evidence, but between the *absolute* amounts of relevant knowledge and of relevant ignorance respectively.
>
> (*TP*, p. 77)

> As the relevant evidence at our disposal increases, the magnitude of the probability of the argument may either decrease or increase, according as the new knowledge strengthens the unfavorable or the favorable evidence; but *something* seems to have increased in either case—we have a more substantial basis upon which to rest our conclusion. I express this by saying that an accession of new evidence increases the weight of an argument. New evidence will sometimes decrease the probability of an argument, but it will always increase its "weight."
>
> (*TP*, p. 77)

Despite the terminological difference between Peirce who used "weight of evidence" in the sense of net weight or balance of argument, and Keynes who, in the passage cited, was using "weight" in the sense of gross weight, Peirce and Keynes were in agreement that balance of argument or probability alone could not characterize all important aspects of evidential appraisal.

That is where the agreement ended. Peirce insisted that numerically determinate belief probability judgements be grounded or supported by knowledge of statistical or physical probability. Such statistical knowledge itself may be imprecise and subject to some kind of random error. It is that random error that Peirce thought could be used to characterize some of the other aspects of evidential appraisal.

In discussing the frequency interpretation of probability, Keynes complained about the narrowness of Venn's approach, which makes no claims about how to ground belief probabilities in statistical probabilities or long-run frequencies.[2]

Keynes did consider a view of the frequency theory different from Venn's, which he attributed tentatively to Karl Pearson (*TP*, p. 109). On this view, as it is according to Peirce's view, knowledge of statistical probability is used to ground judgements of belief probability. Keynes appreciated, as did Peirce, that, according to this version of the frequency view, the problem of choosing a reference class (or kind of trial) becomes critical.

According to both Peirce and the Keynesian reconstruction of the frequency view, the reference class used in direct inference should contain all relevant information about the specific experiment (see Keynes, *TP*, p. 113).

For Peirce, the judgement of relevance is a judgement of statistical relevance—that is, information about statistical or physical probabilities. This is information *X* must know. It is, in this sense, grounded in fact.

According to Keynes, the judgement of relevance is not a judgement of fact or grounded in fact but is a direct judgement about belief probabilities. Indeed, Keynes claimed that when belief probability can be "measured" by a known "truth-frequency," the same result can be obtained by a proper use of his Principle of Indifference—that is, insufficient reason. I believe that Keynes had in mind here a thesis suggestive of Bruno de Finetti's use of symmetry conditions on probability judgements to relate belief probabilities of single states to frequencies in reference classes to which they belong. De Finetti thought that he could replace allusion to physical or statistical probability in all contexts where it seemed useful to replace judgements of statistical probability by judgements of subjective or belief probability. It appears that Keynes did also.

In many cases, information available to the inquirer may not be known to be either statistically irrelevant or statistically relevant, so that Peirce's approach offers no clear guidance as to how to draw conclusions about the outcomes of some kind of trial of a certain kind even when the chances of outcomes of that kind on such a trial are known. We may be convinced that a cab from the city is involved in an accident and 85 per cent of the city's cabs are yellow and the remainder blue without being able to judge via direct inference that the belief probability that the cab in the accident is yellow is 0.85. We may not know whether 85 per cent of cabs in the city involved in accidents are yellow. That is to say, we may not know what percentage of cabs in the city involved in accidents are yellow and, hence, we may be ignorant of the stochastic relevance or irrelevance of the information that the cab in question was involved in an accident. In that case, Peirce himself insisted in 1867 that credal probability judgement goes indeterminate. Ignoring the "base rate" information given is no fallacy.

Keynes's criticism of Venn's frequency view carries over to Peirce. The applicability of Peirce's theory is severely limited. However, Keynes should not have made this objection since he himself conceded that credal probability can be indeterminate.

Keynes could have insisted that indeterminacy would be intolerably widespread if one insisted on Peirce's demands. One needs to be in a position to make moderately determinate judgements of belief probability without grounding in objective or statistical chance—at least for the purpose of assessing relevance.

This is not the place to elaborate on the controversy. I shall say only that I am inclined to think that Peirce's position is overly demanding. This point is not sufficient to undermine the importance of frequentist, statistical, or physical probability as Keynes seems to suggest. It does mean that we cannot restrict the use of determinate or relatively determinate belief probabilities to those grounded in knowledge of statistical probabilities.

Return now to the question of the weight of argument. In his 1878 paper and even more emphatically in 1883 Peirce proposed an account of how to make estimates from data about frequencies without appeal to Bayes' theorem that are probabilistically reliable. In this way, Peirce was able to avoid the use of prior belief probabilities while using data as "input" into the procedure without using it as evidence or premises of his "inference" and violating his strictures on direct inference. The method is essentially the method of confidence interval estimation of Neyman–Pearson (see Levi, 1980, ch. 17).

Such a method of estimation can be associated with a measure of its accuracy determined by the standard deviation, as we noted Peirce proposed to do. Keynes recognized the possibility of using measures of dispersion of a probability distribution as measures of weight of argument. In many contexts the variance of posterior distribution of a certain parameter decreases with an increase of information on which the posterior is based. But as Keynes illustrated this need not be so (see *TP*, pp. 80–2). Examples of dispersion increasing with more evidence can be constructed.

3.5 Is more information always a good thing?

Keynes seemed tentatively to be drawn to the conclusion that the (gross) weight of an argument and its probability are two different properties of it (*TP*, pp. 82–3). But this led to a serious worry. In making decisions, Keynes thought, "we ought to take into account the weight as well as the probability of different expectations." Although Keynes reiterated this thought later on in his *Treatise*, he did not propose anything more than a sketch of a positive account of his own of weight of argument or how it bore on decision making or inquiry.

To cut through the metaphors of balancing and weighing, what is the problem to be considered? Here is an illustrative urn model example.

An urn contains 100 black or white balls. If we invoke Insufficient Reason in the manner condoned by Keynes, each possible constitution of black balls and white balls carries equal prior probability. The prior probability of a selection of one ball turning up a black is 0.5, and that should be the betting rate in evaluating a bet on black.

Suppose a sample with replacement of 100 balls is taken at random and 50 per cent are observed to be black. Bayes's theorem gives back a posterior probability of 0.5 for obtaining a black on the next draw. Betting on obtaining a black on the next draw with equal odds for and against is once more the favorite recommendation.

Keynes thought that the decision maker evaluates the balance of arguments in the same way in both cases considered and, hence, this leads to a similar evaluation of the gambles in terms of expectation. Nonetheless, like many decision makers he would prefer to make the decision when the information about the outcome of sampling is available. Keynes saw this preference in a more general setting than the one just described.

> For in deciding on a course of action, it seems plausible to suppose that we ought to take account of the weight as well as the probability of different expectations. But it is difficult to think of any clear example of this, and I do not feel sure that the theory of "evidential weight" has much practical significance.
>
> Bernoulli's second maxim, that we must take into account all the information we have, amounts to an injunction that we should be guided by the probability of that argument, among those of which we know the premises, of which the evidential weight is the greatest.

> But should not this be re-enforced by a further maxim, according to which we ought to make the weight of our arguments as great as possible by getting all the information we can? It is difficult to see, however, to what point the strengthening of an argument's weight by increasing the evidence ought to be pushed. We may argue that, when our knowledge is slight but capable of increase, the course of action, which will, relative to such knowledge, probably produce the greatest amount of good, will often consist in the acquisition of more knowledge. But there clearly comes a point when it is no longer worth while to spend trouble before acting, in acquisition of further information, and there is no evident principle by which to determine *how far* we ought to carry our maxim of strengthening the weight of our argument. A little reflection will probably convince the reader that this is a very confusing problem.
>
> (Keynes, *TP*, pp. 83–4)

Sometimes choice is peremptory. Sometimes inquirers have the option to delay making a final decision and collecting more information on which to base a decision where exercising this option is costly. And sometimes the option of postponing choice pending the acquisition of additional information is without cost or nearly so. In the latter case, the question then arises: Would delaying choice and obtaining more information always serve the aims of the decision maker better than making a "terminal" choice without further ado? If not, what considerations determine when we should stop inquiring? This is the question about weight of argument that Keynes found confusing.

Suppose that agent X confronting a decision problem faces a choice at some initial time t between (a) choosing among a set of terminal options and (b) postponing that choice to some later time t' and acquiring additional information in the interim. Keynes's vision seems to have been that the new information would be the result of observation or experimentation of some sort. In that case, the option of postponing choice and acquiring new information via observation or experimentation is contemplated in a setting where X does not know what the new information to be acquired will be.

3.6 Keynes and Ramsey (and Savage and Good)

Ramsey explicitly addressed in an unpublished note (see Ramsey, 1990b)[3] Keynes's question understood as a question about the acquisition of new information about data of observation and experiment. Leonard

Jimmie Savage (1954, s. 6.2) and Irving John Good (1967; reprinted in Good, 1983, ch. 17) offered essentially the same account as Ramsey, apparently independently both of each other and of Ramsey.

Ramsey, Savage, and Good all argued that acquiring the new data and then choosing maximizes expected utility provided certain conditions are satisfied. The collection of data should be cost-free. The inquirer *X* should be convinced that he will update *X*'s initial probability judgements using Bayes's theorem via conditionalization. The outcome of experimentation makes a difference as to which option maximizes expected utility. If the same option maximizes expected utility regardless of the outcome of experimentation and observation, the expected utility of acquiring the new data and then choosing is the same as the expected utility of choosing without the benefit of the data.

Let us see how this argument works in a simple case. Suppose agent *X* is told that an urn contains 90 black balls and 10 white balls (H_B), or 10 black balls and 90 white balls (H_W), or 50 black and 50 white balls (H_N). *X* is also told that whether H_B, H_W, or H_N is true depends on the outcome of random process that assigns equal chance of 1/3 to each alternative. As far as *X* is concerned the probabilities of the three hypotheses are equal. *X* is offered three options given in Table 3.1.
X should evaluate the three options equally according to expected utility relative to the initial state of information.

Let *X* now be offered the opportunity to observe the outcome of a random selection of a single ball before making a choice. The probability $P(\text{black}/H_B) = P(\text{white}/H_W) = 0.9$ and $P(\text{white}/H_B) = P(\text{black}/H_W) = 0.1$. $P(\text{black}/H_N) = P(\text{white}/H_N) = 0.5$. Since $P(H_B) = P(H_W) = P(H_N)$ 0.33, by Bayes's theorem $P(H_B/\text{black}) = P(H_W/\text{white}) = 0.6$. $P(H_N/\text{black}) = P(H_N/\text{white}) = 0.33^+$ and $P(H_B/\text{white}) = P(H_W/\text{black}) = 0.07^-$.

Thus, adding the information that the ball drawn is black to *X*'s body of full beliefs and updating via conditionalization and Bayes's theorem lead to a belief probability of 0.6 for H_B. This equals the expected value of A. The expected value of B is 1/15 and for C is 0.1/3. So the result recommends A as the best option. The addition of the information that

Table 3.1 Random process and evaluation of hypotheses

	H_B	H_N	HW
A	1	0	0
B	0	0	1
C	0	1	0

the ball drawn is white determines B as the best option with the same expected value of 0.6. Thus, on the supposition that *X* will maximize expected utility on obtaining the new information whatever it may be, *X* evaluates making the observation and choosing the option that then maximizes expected utility as itself carrying an expected utility of 0.6. This is better than the expected utility 0.33^{+} of choosing any one of A, B, or C without the benefit of the observation.

Can *X* improve *X*'s predicament still further by observing another ball selected from the urn at random? (I shall suppose that the first ball is returned to the urn prior to this second selection.) If the first ball selected is black, the second can be black or white. If black, the posterior for H_B will be boosted even higher (0.76). A will be recommended as before with even higher expectation. If white, the posterior for H_B will be reduced to 0.21. The posterior for H_W will increase to 0.21 and the posterior for H_N will be 0.58. Option C is then optimal. The expected value of obtaining the new information will be 0.75. This is higher than 0.6.

It is demonstrable that collecting new data in this way can never be worse than refusing the new data as long as no extra cost is incurred and we ignore the risk of importing error in acquiring the data.

Appealing to expected utility when deciding whether to obtain more data is not quite the same as appealing to expected utility in choosing between terminal options. Keynes seems quite clear that weight of argument is not relative to any specific decision problem. Moreover, as long as a is not entailed by h, the weight of argument for a can always be increased by strengthening h relevantly. (See the condition "$V\ (a/hh_1) = V\ (a/h)$, unless h_1 is irrelevant, in which case $V\ (a/hh_1) = V\ (a/h)$" in Keynes's *Treatise on Probability*) (*TP*, p. 79).

Nonetheless, Ramsey offered an answer to Keynes's problem. When information is cost free, risk free and relevant to the decision under consideration, it pays to obtain it relative to the aims of the problem at hand.

In spite of its distinguished provenance and the indubitable validity of the argument under the assumptions upon which it is made, the Ramsey argument has severely limited applicability.

Keep in mind that the inquirer *X* is in a context where *X* is deciding whether to perform an experiment and then reach a decision or to take the decision without experiment or observation. In that setting, the inquirer does not, as yet, know whether the experiment and observation will be made and, if it is, what it will be. *X* may be in a position to make determinate probability judgements as to what the outcome of experiment will be conditional on running the experiment or not as the case may be. Ramsey (and Savage and Good) all presupposed that the probabilities would be determinate. If they are not, the import of the argument is quite different.

If we set aside the possibility (which Keynes insisted upon) that probabilities are indeterminate, the calculations upon which the Ramsey–Savage–Good argument is based in our example and in general presuppose that errors of observation are ignored.

Errors of observation would be legitimately ignored if the inquirer *X* were absolutely certain that no such error could arise. However, to suppose that the inquirer rules out the logical possibility that forming the belief that a black (white) is drawn in response to observation when the ball drawn is white (black) is to suppose that *X* is more confident of the testimony of the senses than *X* normally should be. It would not be sound practice to assume in advance of making observations that the observations will be 100 per cent reliable.

Perhaps risk of error should be ignored not because importing false belief is not seriously possible but because it is not important. According to a vulgar form of pragmatism to which Peirce did not subscribe, the inquirer *X* should not attach any particular value to avoiding false belief unless it impacts on the consequences of *X*'s actions relative to the practical goals *X* is committed to realizing. *X* might acknowledge the serious or epistemic possibility of errors of observation and continue to ignore them because they have no impact on *X*'s expectations as to what the practical consequences of *X*'s decisions will be.

Vulgar pragmatists insist that practical considerations always override cognitive goals. So unless risk of error is relevant to promoting or undermining the realization of practical goals, vulgar pragmatists could judge themselves justified in ignoring the possibility of error.

But risk of error can be relevant to the realization of practical goals. Thus, in the case where the issue is to make an observation of the color of one ball drawn from the urn in the case where $p\ (H_B) = 0.33$, if the probability of error is greater than 46 per cent the expected value of making an observation will be less than one-third and, hence, will be disadvantageous. In those cases where the Ramsey argument leads to the result that acquiring new information via observation is neither advantageous nor disadvantageous, taking risk of error into account can make no difference. But where the expected value of the new information is positive, taking risk of error into account can undermine the Ramsey argument even when practical considerations alone are considered.

Taking cognitive values including risk of error seriously could deter the making of observations in cases where there is neither practical advantage nor disadvantage otherwise. Thus, if *X* had already observed a large number of draws from the urn and they were overwhelmingly black, making an additional observation would not make any difference to *X*'s decision to choose option A. But a new observation might incur the risk of a false belief that the ball drawn is black when it is white or white when black. If this risk is slightly greater than the value of the information gained, making the observation would become disadvantageous. Thus, the import of the Ramsey argument is further undermined by insisting on the autonomy of cognitive values.

In the previous section, it was argued that pragmatic justifications for specific inductive inferences should prohibit the overriding of cognitive goals by practical ones. This consideration ought to suffice for the recognition of cognitive values as autonomous dimensions of value and the rejection of vulgar pragmatism and the utilitarianism that so often spawns it.

Ramsey also assumed that the agent *X* is convinced that, upon obtaining the data, *X* will update by conditionalizing on the data to form new probability judgements. But even if *X* is making probability judgements coherently, rationality does not mandate that *X* update probability judgements in this way any more than it mandates changing probabilities by Jeffrey Conditionalization.[4] Rational agents should change probability judgements on acquiring new data via (temporal credal) conditionalization only if they do not revise their confirmational commitments—that is, their commitments as to what probability judgements should be relative to diverse potential states of full belief or evidence (see Levi, 1974; 1980). So even if we restrict discussion to ideally rational agents, *X* must predict that *X* will retain *X*'s current confirmational commitment upon acquiring new information and that such confirmational commitment meets a condition of confirmational conditionalization as a requirement of synchronic rationality. Ideally rational *X* need not do this.

There is another non-trivial presupposition ingredient in the Ramsey argument. *X* must assume prior to acquiring new information that, after obtaining the new information, *X* will choose for the best. At the time t_0 when *X* is contemplating the acquisition of new information, *X* may be in a position to decide whether or not to do so. But *X* is not in control of whether *X* will maximize in the future once the new information is acquired. *X* can predict only at the initial stage t_0 whether *X* will do so or not. And uncertainty may infect this prediction as well.

The reservations registered concerning Ramsey's argument ought not to be taken as a dismissal of its insight. Many of the assumptions tacitly made by advocates of the argument are often reasonably adopted.

The most difficult one, in my judgement, concerns the risk of error. When that risk is sufficiently small, its impact is negligible. A good case can be made for the desirability of acquiring new information when the reservations concerning the assumptions of the Ramsey argument can be overlooked.

Nonetheless, the Ramsey argument does not provide an explication of the notion of weight of argument from evidence *h* to hypothesis *a* where the argument is the judgement of probability that *a* given *h*. Keynes did suggest that perhaps such an assessment of weight would be useful in determining whether the current total evidence is sufficient for terminating investigation and taking whatever practical decision is at issue. But the threshold level might differ depending upon what the practical decision problem is. Keynes was interested in what sort of measure of weight of argument is suitable for the purpose no matter what the threshold might be. The choice of a threshold might depend on the practical goals of the decision problem. Keynes seemed to have been interested in what the threshold is a threshold of. The assessment of weight of argument is in this sense independent of the specific goals of the practical decision problem. It is clear that the Ramsey argument cannot answer the question raised.

3.7 Inductive warrant and weight of argument

Let E be the information about the relative frequency of blacks and whites obtained in a sample with replacement from the urn confronting *X* in the toy example. What constitutes an inductive warrant for adding H_B to K^+_E? K is the inquirer's initial state of full belief or background knowledge. K^+_E is the state of full belief obtained by adding E to K together with all the logical consequences. Remember that, relative to K^+_E, *X* is committed to being certain of the logical consequences of K and E while judging it a serious possibility with positive probability that H_B is false. If there is an inductive warrant for expanding K^+_E by adding H_B, it is a warrant for becoming certain that H_B. After one becomes certain that H_B, there is no point in inquiring further as to the truth or falsity of H_B. As far as that issue is concerned, the weight of argument for or against H_B has reached a maximum.

This line of thinking seems to be consonant with Keynes's own, although his remarks at most gesture in this direction. Keynes explicitly claimed that the weight of argument associated with the probability argument *a*/*h* is always equal to the weight of argument associated with ~*a*/*h*. For an argument is always as near proving or disproving a proposition as it is to disproving or proving its contradictory (*TP*, p. 84).

Thus, according to Keynes the weight of the argument a/h—that is, the weight of the argument *in favor of a* afforded by h—is equal to the weight of the argument $\sim a/h$—that is, the weight of the argument *against a* afforded by h. This suggests that lurking behind weight of argument are two dual notions of positive warrant or support $b(a/h)$ for a by h and negative warrant or support d(a/h) for a by h. Clearly, $d(\sim a/h) = b(a/h)$.

The idea is that given the weight of the argument from h to a, that weight can support or "prove" to some degree that a or it can disprove it or it can do neither. When it supports a, it disproves $\sim a$.

Given the pair of values $d(a/h)$ and $d(\sim a/h)$ [i.e., $b(\sim a/h)$ and $b(a/h)$], the weight of argument for or against a given h is the maximum value in the pair.

What happens in cases where the weight of the argument in favor of a equals the weight of the argument in favor of $\sim a$? Keynes introduced the idea of "nearness" to proving or to disproving and, hence, of positive and negative warrant. But he did not seem to consider explicitly the case where the negative and the positive warrants for proposition a relative to h are equal except insofar as it can be teased out of applications of Insufficient Reason to probability judgements. If the probability of a given h equals the probability of $\sim a/h$, the weight of argument in favor of a and in favor of $\sim a$ relative to h ought to be the same (see *TP*, pp. 79–80).

Let h and the background knowledge entail that exactly one of a, b, and c is true and that the conditions for applying the Principle of Indifference obtain. Each of the propositions gets probability 1/3. According to Keynes (*TP*, pp. 78–9) the weights of the arguments from h to each of the three alternatives are equal. So are the weights of the arguments for each of the negations.

The circumstances just envisaged are precisely of the sort where the argument inferring a from h is "as near" proving a from h as the argument inferring $\sim a$ from h is near to disproving a from h.

Consequently, we cannot take a measure of proximity of the inference from h to a as proof of a to be the probability 1/3 of a (see Keynes, *TP*, p. 80). The proximity of the inference from h to $\sim a$ would then be 2/3. But the one inference is supposed to be as close to proof as the other is. We could take the proximity of the inference from h to a to a proof and the proximity of the inference from h to $\sim a$ to a proof to be ½ or we could take it to be any non-negative value we like including 0.

The important point is that, whatever value we take, the proximity of a given h to proof and the proximity of $\sim a$ given h to proof are both at a minimum. An increase in the proximity of a given h to proof corresponds to a decrease in the proximity of $\sim a$ given h to proof (that is, to an increase in the proximity of $\sim a$ given h to disproof) and vice versa. Keynes's own appeal to the notion of proximity to proof and disproof as characterizing weight of argument hints at this much.

This suggests that 0 is a convenient value to adopt for the minimum. And, of course, Ramsey's argument is most compelling precisely when the data h provide minimum proof for a and for $\sim h$.

Once more, by Keynes's own principles, the weight of argument for $a \vee b$ given h ought to be equal to the weight of argument for $\sim a \vee \sim b$ given h. The proximity to full proof of the former ought to equal proximity to full disproof of the latter and vice versa.

Here it seems plausible to take a step beyond Keynes's explicit discussion. The proximity of $a \vee b$ given h to full proof ought to be the minimum of the proximity of a given h to full proof and of b given h to full proof. Likewise the proximity of $\sim a \vee \sim b$ given h to full proof ought to be equal to the minimum of a given h to full disproof and of b given h to full disproof. No argument can come closer to proving $a \vee b$ from h than the argument from h to the conjunct that is least close to full proof.

These observations are based on *very* slender threads of textual evidence in Keynes. The reasoning I have sketched is based, nonetheless, on suggestions that are found in Keynes himself when focusing on the evaluation of the weights of different hypotheses given fixed evidence h.

This reasoning points to the idea that the b-functions or d-functions used to define weight of argument given fixed evidence h exhibit the properties of George L. S. Shackle's measures for potential surprise or disbelief and the dual notion of degree of belief (see Shackle, 1952; 1961). The formal properties of Shackle-type belief and disbelief parallel those I have sketched for proximity to proof and disproof. One can use the one measure or the other to represent weight of argument.

Space does not permit illustration of this understanding of weight of argument. Nonetheless, I suggest that the notion of weight of argument that Keynes was seeking might be interpreted by reference to the specificity of conclusions warranted with a Shackle degree of confidence relative to K (see Levi, 1967, 1984, 1996 and 2001 for elaboration).

Notes

1. I am proposing here to reconstruct Keynes's view as recommending that rational agents endorse a rule for adopting states of probability judgement relative to diverse potential states of full belief or certainty, the weakest confirmational commitment in the sense of Levi (1980) allowed by probability logic. Ramsey seemed more inclined to assume that the degrees of belief

or subjective probability would or should be always cardinally measurable, whereas Keynes did not. Like many of his frequentist opponents, Keynes thought that, in the absence of logically compelling determinations, degrees of belief or subjective probability ought to be indeterminate. Ramsey refused to countenance this and deployed some unimpressive and question-begging arguments against Keynes (see Ramsey, 1990a, pp. 56–7.) Ramsey's insight that probability logic could not be both a logic of truth and a logic of consistency registers a more profound objection to Keynes's approach. But that does not touch the reconstruction I am offering.

2. Venn (1888, ch. 9) did make some remarks pertinent to the selection of reference classes and direct inference. But Keynes was essentially correct. Venn thought that the reference class used in direct inference is a practical matter. Keynes complained that Venn was too narrow in restricting probability only to statistical probability. In his review of the 1866 edition of Venn's book, Peirce (1867) complained for the same reason that Venn is too much of a conceptualist because belief probabilities derived from information about frequencies according to Venn are based on appeals to reference classes without objective grounding. Peirce complained especially about Chapter 17 of Venn's book on extraordinary stories.
3. Ramsey did not explicitly mention Keynes's discussion of weight of argument; but there seems little doubt that the note was written in response to Keynes's discussion.
4. B. Skyrms (1990) discusses the Ramsey note (Ramsey, 1990b). Skyrms is anxious to suggest that Ramsey is a forerunner of "Jeffrey updating" or "probability kinematics" and cites this note as an intimation of it along with some notes taken on a paper of Donkin's from the 1850s. I myself can detect no such intimation and suspect that Ramsey may not have even entertained the idea.

References

Good, I. J. (1981), 'An Error by Peirce Concerning Weight of Evidence,' *Journal of Statistical Computation and Simulation*, 13, pp. 155–7.

——— (1983), *Good Thinking*: *The Foundations of Probability and Its Applications*, Minneapolis: Minnesota, University of Minnesota Press.

Keynes, J.M. (1973 [1921]), *A Treatise on Probability*, vol. VIII of *The Collected Writings of John Maynard Keynes*, London: Macmillan.

Levi, I. (1967), *Gambling with Truth*, New York: Knopf, Cambridge, Mass.: MIT Press. Paperback edn 1973.

——— (1974), 'On Indeterminate Probabilities', *Journal of Philosophy*, 71 (July), pp. 391–418.

——— (1980), *The Enterprise of Knowledge*, Cambridge, Mass.: MIT Press. Paperback edn. 1983.

——— (1984), *Decisions and Revisions*, Cambridge: Cambridge University Press.

——— (1996), *For the Sake of the Argument*, Cambridge: Cambridge University Press.

——— (2001), 'Inductive Expansion and Nonmonotonic Reasoning,' in M. A. Williams and H. Rott (eds), *Frontiers in Belief Revision*, Dordrecht: Kluwer, pp. 7–56.

Peirce, C. S. (1867), 'Review of 1866 Edition of *The Logic of Chance* by J. Venn,' in E. C. Moore (ed.), *Writings of Charles S. Peirce*, 2 Bloomington, Ind., Indiana University Press (1984), pp. 98–102.

——— (1878), 'Probability of Induction,' in C. J. W. Kloessel, (ed.), *Writings of Charles S. Peirce*, 3, Bloomington, Ind.: Indiana University Press (1986), pp. 290–305.

——— (1883), 'A Theory of Probable Inference', in C. J. W. Kloessel, (ed.), *Writings of Charles S. Peirce*, 4, Bloomington, Ind.: Indiana University Press (1986), 408–50.

Ramsey, F. P. (1990a), *Philosophical Papers*, Cambridge: Cambridge University Press.

——— (1990b), 'Weight or Value of Knowledge,' *British Journal for the Philosophy of Science* 41, pp. 1–4, with a preamble by N.-E. Sahlin.

Savage, L. J. (1954), *The Foundations of Statistics*, New York: Wiley; 2nd revised edn 1972, New York: Dover.

Shackle, G. L. S. (1952), *Expectation in Economics*, 2nd edn (1st edn 1949), Cambridge: Cambridge University Press.

——— (1961), *Decision, Order and Time*, Cambridge: Cambridge University Press.

Skyrms, B. (1990), *The Dynamics of Rational Deliberation*, Cambridge. Mass.: Harvard University Press.

Venn, J. (1888), *The Logic of Chance*, 4th edition.