

Good Guesses as Accuracy-Specificity Tradeoffs

Mattias Skipper

Penultimate draft, forthcoming in *Philosophical Studies*

Abstract: Guessing is a familiar activity, one we engage in when we are uncertain of the answer to a question under discussion. It is also an activity that lends itself to normative evaluation: some guesses are *better* than others. The question that interests me here is what makes for a good guess. In recent work, Dorst and Mandelkern have argued that good guesses are distinguished from bad ones by how well they optimize a tradeoff between *accuracy* and *specificity*. Here, I argue that Dorst and Mandelkern’s implementation of this idea fails to satisfy some plausible constraints on good guesses, and I develop an alternative implementation that satisfies the relevant constraints. The result is a new account of good guesses which retains the positive aspects of Dorst and Mandelkern’s proposal, but without the drawbacks.

1. Introduction

A five-horse race is about to start. You don’t know which horse is going to win, but here are their respective winning chances:

Ajax	Benji	Cody	Dusty	Ember
60%	18%	14%	6%	2%

Take a guess: Who do you think is going to win?

There are many guesses you *could* make. But some guesses are clearly better than others. While ‘Ajax’ and ‘Ajax, Benji, or Cody’ both seem like fine guesses, ‘Dusty’ and ‘Cody or Ember’ both seem like terrible ones. What explains the difference? What makes for a good guess? That is the question I want to address in this paper.

There are at least two reasons for being interested in this question. First, as we will see, good guesses turn out to exhibit some striking patterns that call out for explanation in their own right. Second, the notion of guessing has itself been put to explanatory use in recent attempts to shed light on the cognitive attitudes of *thinking* and *believing* (Holguín 2022; Hawthorne et al. 2016). I won’t here take a stance on whether these deployments of the notion of guessing are apt. But if the notion of guessing is to do substantive work in epistemology or elsewhere, it would at least be good to have a basic understanding of what makes for a good guess.

Perhaps the most systematic attempt to develop an account of good guesses comes from a recent paper by Dorst and Mandelkern (forthcoming), who argue that good guesses are distinguished from bad ones by how well they optimize a tradeoff between *accuracy* and *specificity*. As we will see, there is much to recommend their proposal. Their central idea, that good guesses optimize a tradeoff between accuracy and specificity, seems to me a promising

one. Nonetheless, I will argue that their implementation of this idea fails to satisfy some plausible constraints on good guesses (§§2-3), and I will develop an alternative implementation that satisfies the relevant constraints (§4). The result will be a new account of good guesses which retains the positive aspects of Dorst and Mandelkern’s proposal, but without the drawbacks.

2. Dorst and Mandelkern on Good Guesses

Let us start by asking: what do we *want* from our guesses? We obviously want them to be *correct*—we want to make *accurate* guesses. But accuracy cannot be all we care about. After all, it is easy to ensure that your guess is correct: just guess the disjunction of every possible outcome! In the horse race scenario, this would amount to guessing ‘Some horse will win.’ That’s clearly not a good guess. We also want to *take a stand* on things—to make a *specific* guess, one that narrows down the set of possibilities under consideration.¹ These goals directly compete: the more specific your guess is, the less likely it is to be accurate; and the less specific it is, the more likely it is to be accurate. Thus, we face a tradeoff: we must strike a balance between making guesses that are likely to be accurate, but not too unspecific; specific, but not too unlikely to be accurate. In making this tradeoff, some agents might be more or less “risk-averse” than others. Someone who is risk-averse (in this sense) will place more weight on accuracy than specificity, whereas someone who is risk-taking will place more weight on specificity than accuracy. And what constitutes a good guess depends on how risk-averse or risk-taking you are: a good guess is one that optimizes your preferred accuracy-specificity tradeoff.

This summarizes the basic idea behind Dorst and Mandelkern’s proposal. My concern here is how this idea is best made precise, so I’ll begin by taking a closer look at the details of Dorst and Mandelkern’s account. We begin with a set of possibilities, \mathcal{W} , which we think of as the set of scenarios or “possible worlds” that are compatible with what the agent is certain of. A *question*, Q , is a partition of \mathcal{W} : a set of mutually exclusive and exhaustive subsets of \mathcal{W} .² The cells of the partition are the *complete answers* to Q . The *size* of Q , denoted $|Q|$, is the number of complete answers to Q . The agent’s credences are represented by a probability function, P , which is defined over the subsets of \mathcal{W} . Since the agent is certain of \mathcal{W} , and since the complete answers form a partition of \mathcal{W} , the agent is certain that exactly one of the complete answers is correct.

To illustrate these definitions, let us they apply them to the horse race example. The question under discussion is: “Which horse is going to win?” There are five complete

¹ Dorst and Mandelkern use the term “informativity” to refer to what I’m calling “specificity” here. Nothing of importance turns on this choice of terminology, but I prefer the term “specificity” because the term “informativity” is sometimes used by information theorists for a different purpose (roughly, to capture how much an agent would learn from coming to know with certainty that a given proposition is true).

² This way of modelling questions, as partitions of logical space, draws on classic work on the semantics of questions—see, e.g., Hamblin (1973) and Groenendijk and Stokhof (1984).

answers: $\{Ajax, Benji, Cody, Dusty, Ember\}$. The size of the question is: $|Q| = 5$. And the probability distribution over the complete answers is: $\{Ajax: 60\%, Benji: 18\%, Cody: 14\%, Dusty: 6\%, Ember: 2\%\}$.

That is the basic setup. Next, we need to make precise the idea that good guesses optimize an agent's preferred accuracy-specificity tradeoff. To this end, Dorst and Mandelkern assign to each answer, p , an *answer-value*, $V(p)$, which represents how good it would be to guess p in response to the question under discussion. Since we want our guesses to be accurate, the answer-value of p depends, in part, on whether p is true. So, whenever you're uncertain about whether p is true, you will be uncertain of p 's answer-value. Still, you can form an *expectation* of p 's answer-value. Let $V^+(p)$ be p 's answer-value *if true*, and let $V^-(p)$ be p 's answer-value *if false*. The expected answer-value of p is then given by the following weighted average:

$$E(p) = P(p) \cdot V^+(p) + P(\sim p) \cdot V^-(p).$$

According to Dorst and Mandelkern, this is the quantity one should aim to maximize when forming one's guess. That is, Dorst and Mandelkern propose the following norm of guessing:

Guessing as Maximizing: A guess is permissible iff no other guess has a higher expected answer-value.

Before we can derive any predictions from this norm, we need to say more about what determines the answer-value of a guess. Recall that the answer-value of a guess is supposed to depend on two factors: its accuracy and its specificity. To capture the accuracy-component, Dorst and Mandelkern require that V be *truth-directed*:

Truth-directedness: $V^+(p) > V^-(p)$.

This is just to say that true guesses are better than false ones. In addition, Dorst and Mandelkern make the simplifying assumption that false guesses never have any positive or negative answer-value (i.e., $V^-(p) = 0$), and hence that true guesses always have a positive answer-value (i.e., $V^+(p) > 0$).³ This allows us to simplify the expression for the expected answer-value of a guess as follows:

$$E(p) = P(p) \cdot V^+(p).$$

Since $V^-(p)$ has dropped out of the equation, I will henceforth refer to $V^+(p)$ simply as "the answer-value of p ."

³ This assumption may not be entirely innocuous. For example, it's natural to wonder whether it isn't better to make a specific false guess than an unspecific false guess, just as it is better to make a specific true guess than an unspecific true guess. For the purposes of this paper, however, I'm happy to proceed on the assumption that false guesses always have zero value or disvalue.

So much for the accuracy-component. What about the specificity-component? Just as true guesses are better than false ones, specific guesses are better than unspecific ones. So, V^+ should not only be truth-directed, but also *specificity-directed*.

Specificity-directedness: $V^+(p)$ is an increasing function of p 's specificity.

To make this requirement precise, we obviously need to say more about what determines the specificity of a guess—we need a *measure* of specificity. What should such a measure look like? We have said that the specificity of a guess is supposed to tell us something about how far the guess “narrows down” the set of possibilities under consideration. And how far a guess can narrow down the set of possibilities under consideration depends on the size of the question under consideration: if the size of the question is quite large—i.e., if the question divides up the set of possibilities in a quite fine-grained way—there is a potential to rule out a large number of possibilities, whereas if the size of the question is quite small—i.e., if the question divides up the set of possibilities in a quite coarse-grained way—there is only a potential to rule out a small number of possibilities. To capture this dependency, Dorst and Mandelkern propose to measure the specificity of a guess in terms of the proportion of complete answers it rules out:

Dorst-Mandelkern Specificity: The *Dorst-Mandelkern specificity* of a guess, p , in response to a question, Q , is given by:

$$S_{\text{DM}}(p) = \frac{|\mathcal{Q}| - c_p}{|\mathcal{Q}|},$$

where c_p is the number of complete guesses, which are compatible with p :
 $c_p = |\{q \in \mathcal{Q}: p \cap q \neq \emptyset\}|$.

If, for example, the question is “Which horse is going to win?,” then $S_{\text{DM}}(\textit{Ajax}) = 4/5$, $S_{\text{DM}}(\textit{Benji or Dusty}) = 3/5$, $S_{\text{DM}}(\textit{Some horse will win}) = 0$, and so on.

Note that $S_{\text{DM}}(p)$ is a decreasing function of c_p , and an increasing function of $|\mathcal{Q}|$: when $c_p = |\mathcal{Q}|$ (i.e., when p includes every complete answer), $S_{\text{DM}}(p) = 0$, and when $c_p = 1$ (i.e., when p is itself a complete answer), $S_{\text{DM}}(p) = (|\mathcal{Q}| - 1)/|\mathcal{Q}|$, which is a quantity that approaches 1 as $|\mathcal{Q}|$ approaches infinity. Hence, Dorst and Mandelkern’s specificity measure faithfully captures the idea that a specific guess is one that narrows down the space of possibilities under consideration, and that how far a guess can narrow down the space of possibilities depends on the size of the question under consideration. Later on, I will raise some concerns about their measure, and I will propose an alternative measure of specificity instead. But before I turn to the critical part of the paper, I want to introduce the rest of Dorst and Mandelkern’s account.

The final task is to capture the idea that different agents might have different preferred accuracy-specificity tradeoffs. To this end, Dorst and Mandelkern introduce a parameter, J ,

which is supposed to represent the degree to which an agent values specificity relative to accuracy: the higher the value of J , the more the agent values specificity relative to accuracy.⁴ Thus, the answer-value of p should be an increasing function of J , just as it should be an increasing function of p 's specificity. From a purely mathematical point of view, there are many different measures of answer-value that satisfy this constraint, but Dorst and Mandelkern opt for the following measure (for reasons I will review shortly):

Dorst-Mandelkern Answer-value: The *Dorst-Mandelkern answer-value* of a guess, p , in response to a question, Q , is given by:

$$V_{\text{DM}}^+(p) = J^{S_{\text{DM}}(p)} = J^{(|Q| - \varphi_p)/|Q|},$$

where $J \geq 1$.

So, for example, if the question is “Which horse is going to win?,” $V_{\text{DM}}^+(\textit{Ajax}) = J^{1/5}$, $V_{\text{DM}}^+(\textit{Benji or Dusty}) = J^{2/5}$, $V_{\text{DM}}^+(\textit{Some horse will win}) = J$, and so on. Note that $V_{\text{DM}}^+(p)$ is indeed an increasing function of both J and p 's specificity (except when $J = 1$, in which case $V_{\text{DM}}^+(p) = 1$ regardless of the value of $S_{\text{DM}}(p)$, which reflects the fact that the agent places no weight at all on specificity).

Given this measure of answer-value, the *expected* answer-value of a guess becomes:

$$E_{\text{DM}}(p) = P(p) \cdot J^{S_{\text{DM}}(p)} = P(p) \cdot J^{(|Q| - \varphi_p)/|Q|}.$$

What Dorst and Mandelkern's account says, then, is that p is a permissible guess iff there is no other guess, q , such that $E_{\text{DM}}(q) > E_{\text{DM}}(p)$.

To illustrate how their account works, consider again the horse race example:

Ajax	Benji	Cody	Dusty	Ember
60%	18%	14%	6%	2%

Which guess has the highest expected Dorst-Mandelkern answer-value? It depends on the J -value: given a low enough J -value, the best guess will simply be the most probable answer, which is always going to be the entire disjunction (“Some horse will win”). Given a high enough J -value, the best guess will be the most probable complete answer (“Ajax”). And for intermediate J -values, the best guess will be either ‘Ajax or Benji’, ‘Ajax, Benji, or Cody’, or ‘Ajax, Benji, Cody, or Dusty’, in order of decreasing J -value (as illustrated by Figure 1).

⁴ \mathcal{J} alludes to William James (1896) who famously discussed a similar tradeoff between the epistemic goals of *believing what is true* and *not believing what is false*.

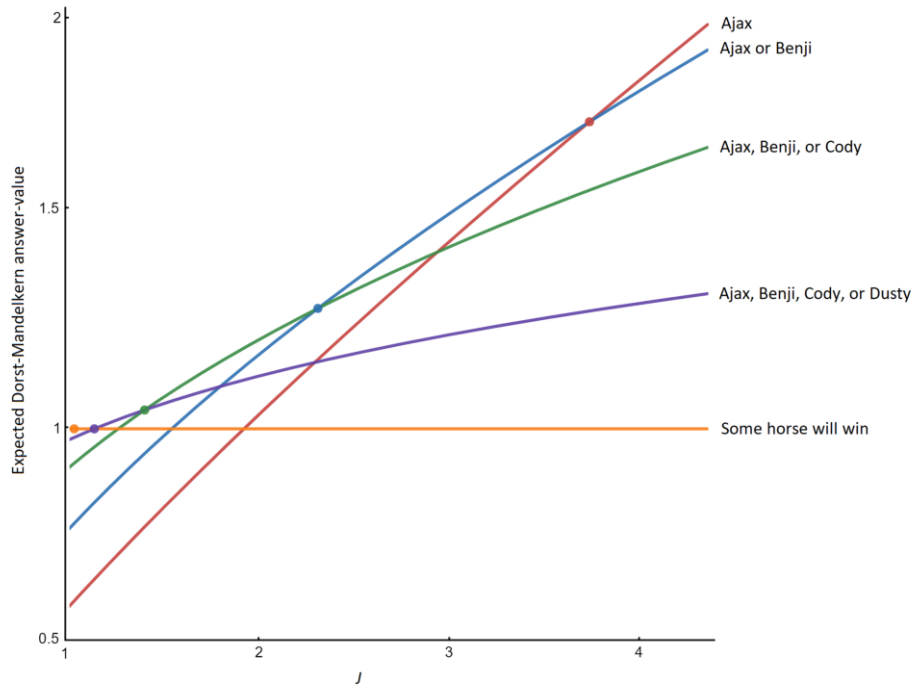


Figure 1: Expected Dorst-Mandelkern answer-value of various horse race guesses, depending on the J -value.

Why think this is the right way of implementing the idea that good guesses optimize an agent’s preferred accuracy-specificity tradeoff? Dorst and Mandelkern helpfully address this question by considering various putative constraints on good guesses, which their account satisfies. Here is one such constraint (which Dorst and Mandelkern attribute to Holguín 2022, §10):

Filtering: A permissible guess is always *filtered*: if it includes a complete answer, p , it must include all complete answers that are more probable than p .

The motivation behind Filtering is straightforward: if p isn’t filtered, there must exist two complete answers, a and a' , such that (i) a is less probable than a' , and (ii) p includes a , but not a' . We can then define an alternative guess, p' , which is identical to p except that it includes a' instead of a . Given this, p and p' are equally specific (they include the same number of complete answers), but p' is more probable than p , which means that p' is a better guess than p regardless of one’s preferred accuracy-specificity tradeoff.

It’s worth noting that Filtering can be equivalently formulated as follows:⁵

⁵ *Proof.* If p is accuracy-dominated by p' , there must exist two complete answers, a and a' , such that (i) p includes a , but not a' , (ii) p' includes a' , but not a , and (iii) $P(a') > P(a)$. But then p isn’t filtered. And we have already seen that unfiltered guesses are accuracy-dominated. So, a guess is filtered iff it isn’t accuracy-dominated.

No Accuracy-Dominance: A permissible guess is never *accuracy-dominated*, where p accuracy-dominates p' iff $c_p \leq c_{p'}$ and $P(p) > P(p')$.

This formulation of Filtering makes it even clearer why it must hold: it follows immediately from the idea that accurate guesses are, all else being equal, better than inaccurate ones. Furthermore, it makes salient a different constraint, which Dorst and Mandelkern do not explicitly discuss, but which their account satisfies:

No Specificity-Dominance: A permissible guess is never *specificity-dominated*, where p specificity-dominates p' iff $c_p < c_{p'}$ and $P(p) \geq P(p')$.

Like No Accuracy-Dominance, this constraint is easy to motivate: it follows immediately from the idea that specific guesses are, all else being equal, better than unspecific ones. This also means that neither constraint offers much help in deciding between different measures of answer-value: all they tell us is that our measure of answer-value must be truth-directed and specificity-directed.

However, Dorst and Mandelkern consider an additional constraint (also from Holguín 2022, §10) which places considerably more stringent demands on a measure of answer-value:

Optionality: For any question, Q , and any k : $1 \leq k \leq |Q|$, some accuracy-specificity tradeoff renders it permissible to guess the disjunction of exactly k complete answers.

We can think of this constraint as saying that it is “up to the agent” to decide how many complete answers they want to include in their guess. For example, in the horse race scenario, you may include anywhere from one to five horses in your guess, depending on your preferred accuracy-specificity tradeoff. As we have seen, this is precisely what Dorst and Mandelkern’s account predicts (see again Figure 1). That is no accident: as Dorst and Mandelkern show, their account satisfies Optionality in full generality.

Why think that an account of good guesses should satisfy Optionality? Dorst and Mandelkern’s main motivation—which can also be found in Holguín (2022, §10)—comes from reflecting intuitively on the kinds of answers one might sensibly give in response to a question like “Which horse is going to win?”. It clearly makes sense to guess the outright favorite ‘Ajax’. But if you prefer to take a less opinionated stance, it could also make sense to make a disjunctive guess like ‘Ajax or Benji’ or ‘not Ember’. And if you really don’t want to stick your neck out, it could even make sense to answer ‘Some horse will win’ by reference to the fact that all five horses have a non-zero chance of winning.

It seems, then, that the norms of guessing are quite permissive: they don’t force us to include any particular number of complete answers in our guess. This is what Optionality is supposed to capture. Eventually, I will offer some reasons to doubt that something as strong as Optionality holds without exception. But at this point in the dialectic, I just want to acknowledge that Optionality has a good deal of intuitive appeal, which offers at least some *prima facie* support for Dorst and Mandelkern’s account.

3. Problems for Dorst and Mandelkern’s Account

I now want to raise some concerns about Dorst and Mandelkern’s proposal. The concerns center around two putative constraints on good guesses, which I think an account of good guesses should satisfy, and which Dorst and Mandelkern’s account turns out not to satisfy. Below I introduce each constraint, explain why I find it attractive, and show how Dorst and Mandelkern’s account violates it. This will then lead to the positive proposal of the paper, which I will present in §4.

3.1 Neutrality

To illustrate the first concern, let us begin by considering a simple coin flip:

Heads	Tails
50%	50%

Compare the guesses ‘Heads’ (H) and ‘Heads or Tails’ ($H \vee T$). Which is better? It clearly depends on your preferred accuracy-specificity tradeoff: if you value accuracy to a high enough degree, $H \vee T$ will be the better guess, and if you value specificity to a high enough degree, H will be the better guess. In other words, for low enough J -values, $H \vee T$ will be the better guess, and for high enough J -values, H will be the better guess. Given this, we should expect that if we set J at just the right intermediate value, H and $H \vee T$ will be tied: neither guess will be better than the other. This is exactly what Dorst and Mandelkern’s account predicts: H and $H \vee T$ have the same expected Dorst-Mandelkern answer-value iff $J = 4$.⁶ So far, no problem.

But now suppose we replace the coin with a six-sided die:

1	2	3	4	5	6
1/6	1/6	1/6	1/6	1/6	1/6

Let us again compare one of the maximally specific guesses, say ‘1’, to the entire disjunction ‘ $1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6$ ’. Which is better? When $J = 4$, it turns out that ‘1’ has a *lower* expected Dorst-Mandelkern answer-value than ‘ $1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6$ ’.⁷ In other words, Dorst and Mandelkern’s account predicts that when you are indifferent between H and $H \vee T$, you won’t be indifferent between ‘1’ and ‘ $1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6$ ’. And you won’t be indifferent between the various intermediate guesses (‘ $1 \vee 2$ ’, ‘ $1 \vee 2 \vee 3$ ’, and so on) either: when you are indifferent

⁶ Calculation: $E_{DM}(p) = E_{DM}(p \vee \sim p) \iff$
 $.5 \cdot J^{2-1/2} = 1 \cdot J^{2-2/2} \iff$
 $.5 \cdot J^{1/2} = 1 \iff$
 $J = 4.$

⁷ Calculation: $E_{DM}(1) = (1/6) \cdot J^{6-1/6} = .53 < 1 = (6/6) \cdot J^{6-6/6} = E_{DM}(1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6).$

between ‘ H ’ and ‘ $H \vee T$ ’, you will consider ‘ $1 \vee 2 \vee 3 \vee 4$ ’ to be the uniquely best guess in the die roll case (see Figure 2 for an illustration).⁸

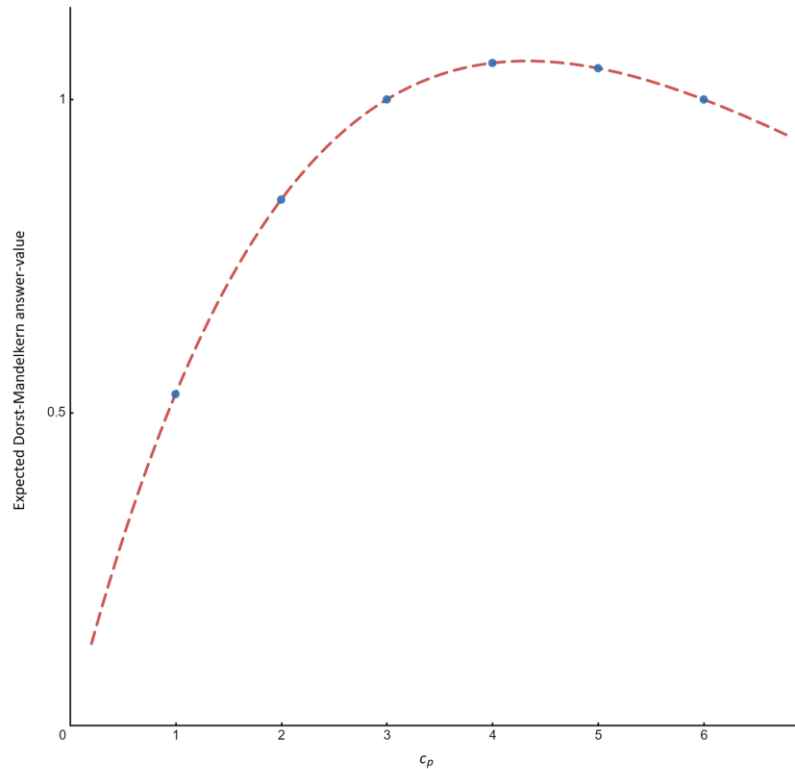


Figure 2: Expected Dorst-Mandelkern answer-value of various die roll guesses (setting $J = 4$).

This strikes me as an odd result. There doesn’t seem to be anything unnatural about being indifferent between ‘ H ’ and ‘ $H \vee T$ ’ while also being indifferent between ‘1’, ‘ $1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6$ ’, and the various intermediate guesses in the die roll case (‘ $1 \vee 2$ ’, ‘ $1 \vee 2 \vee 3$ ’, and so on). On the contrary, it seems possible to justify such a guessing pattern by reference to the fact that when the probability distribution over the complete answers is flat, there is a sense in which you have *nothing to go on* when forming your guess—your guess won’t be *informed* by anything. Of course, you might still prefer some guesses over others on the grounds that you are risk-averse or risk-seeking. But it’s natural to think that there should be a “risk-neutral” perspective from which all guesses look equally good or bad when the probability distribution is flat.

If this isn’t already clear enough, the following decision-theoretic analogy may help drive home the point. Suppose you are offered the following bets by a fair bookie who assigns equal probability to ‘Heads’ and ‘Tails’ respectively:

⁸ Of course, since the probability distribution is uniform, ‘ $1 \vee 2 \vee 3 \vee 4$ ’ will be tied with any other four-cell guess (‘ $2 \vee 3 \vee 4 \vee 5$ ’, ‘ $1 \vee 3 \vee 5 \vee 6$ ’, etc.), so the more precise thing to say is that any four-cell guess will have a higher expected Dorst-Mandelkern answer-value than any guess which doesn’t include exactly four complete answers.

Bet	Probability	Payoff
H	$\frac{1}{2}$	\$2
$H \vee T$	1	\$1

If you had to take one of these bets, which one would you pick? It clearly depends on how risk-averse or risk-taking you are: if you are risk-averse, you will prefer to bet on ' $H \vee T$ ', and if you are risk-taking, you will prefer to bet on ' H '. But if you are risk-neutral, you will be indifferent between the bets, since each bet has an expected value of \$1.

The same goes for the following set of bets:

Bet	Probability	Payoff
1	$\frac{1}{6}$	\$6
$1 \vee 2$	$\frac{2}{6}$	\$6/2
$1 \vee 2 \vee 3$	$\frac{3}{6}$	\$6/3
$1 \vee 2 \vee 3 \vee 4$	$\frac{4}{6}$	\$6/4
$1 \vee 2 \vee 3 \vee 4 \vee 5$	$\frac{5}{6}$	\$6/5
$1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6$	$\frac{6}{6}$	\$6/6

Again, which bet you will prefer depends on how risk-averse or risk-taking you are: if you are sufficiently risk-averse, you will prefer to bet on the entire disjunction ' $1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6$ ', and if you are sufficiently risk-taking, you will prefer to bet on '1'. But if you are risk-neutral, you will be indifferent between the bets, since each bet has an expected value of \$1.

It is natural, then, to suppose that there exists a risk-neutral perspective from which all of the above bets look equally good. And although I don't want to make any general claims about the relationship between guessing and betting, I submit that guessing is similar to betting in this respect: it is natural to suppose that there exists a risk-neutral perspective from which all guesses look equally good when the probability distribution is flat.

If this is right, then an account of good guesses should satisfy the following constraint:

Neutrality: Some accuracy-specificity tradeoff is "risk-neutral" in the sense that, for any question, Q , if the probability distribution over the complete answers to Q is flat (that is, if each of the complete answers has a probability of $1/|Q|$), all guesses, whether complete or disjunctive, have the same expected answer-value.

Yet, as we have seen, Dorst and Mandelkern's account fails to satisfy Neutrality. This is the first concern I wanted to raise.

Before we move on to the second concern, it's worth pausing to consider whether something stronger than Neutrality might hold. We have just said that there exists a neutral accuracy-specificity tradeoff when the probability distribution is flat. Given this, it is natural

to wonder whether there might *always* exist a neutral accuracy-specificity tradeoff, even when the probability distribution isn't flat. Might there exist an accuracy-specificity tradeoff relative to which all guesses are equally good, regardless of what the probability distribution looks like?

I think the answer to this question is negative. Consider the following version of the horse race example:

Ajax	Benji	Cody	Dusty	Ember
50%	40%	5%	3%	2%

Suppose you are indifferent between 'Ajax' and 'Ajax or Benji'. What does this tell us about your preferred accuracy-specificity tradeoff? Presumably, it tells us that you are not vastly more concerned with accuracy than specificity, or else you would have preferred 'Ajax or Benji' to 'Ajax'. Conversely, it tells us that you are not vastly more concerned with specificity than accuracy, or you would have preferred 'Ajax' to 'Ajax or Benji'. Instead, you prefer a fairly balanced tradeoff between accuracy and specificity. We can then ask: given your preferred accuracy-specificity tradeoff, how will you rate 'Ajax, Benji or Cody' in comparison to 'Ajax or Benji'? Presumably, you won't think that the very modest boost in accuracy (5 percentage points) makes up for the relatively significant loss of specificity.⁹ In other words, you will consider 'Ajax, Benji or Cody' to be a worse guess than 'Ajax or Benji', and hence a worse guess than 'Ajax' as well. So if you are indifferent between 'Ajax' and 'Ajax or Benji', it looks like you won't be indifferent between 'Ajax', 'Ajax or Benji' and 'Ajax, Benji or Cody'.¹⁰

Of course, this doesn't show that the probability distribution *must* be flat for there to exist an accuracy-specificity tradeoff relative to which all answers are equally good. But it shows that we shouldn't in general expect such a neutral accuracy-specificity tradeoff to exist unless the probability distribution is flat.

3.2 Independence of Irrelevant Alternatives (for Guessing)

The second concern I want to raise for Dorst and Mandelkern's proposal can be illustrated with another simple example. Suppose a golf tournament is about to start. As so often before, Tiger Woods is the heavy favorite, followed by Phil Mickelson at a distant second:

Woods	Mickelson	Other
98%	1%	1%

⁹ If this isn't already clear, we can always lower Cody's winning chances by adding more horses to the race.

¹⁰ As the reader can verify for herself, Dorst and Mandelkern's account and the account proposed in this paper both have this implication.

Compare the guesses ‘Woods’ and ‘Woods or Mickelson’. Which is better? It’s clear that ‘Woods’ is a good deal more specific than ‘Woods or Mickelson’ while being only slightly less probable. So if you place at least a bit of weight on specificity, you will presumably consider ‘Woods’ to be a better guess than ‘Woods or Mickelson’. This is precisely what Dorst and Mandelkern’s account predicts: ‘Woods’ has a higher expected Dorst-Mandelkern answer-value than ‘Woods or Mickelson’ iff $J > 1.03$.¹¹ So far, no problem.

But now let us suppose that instead of lumping all of the competitors to Woods and Mickelson together into one category, ‘Other’, we separate them out (let us say that there are n such players):

Woods	Mickelson	P_1	P_2	...	P_n
98%	1%	$1\%/n$	$1\%/n$		$1\%/n$

Effectively, we have made the option ‘Other’ more fine-grained. Should such fine-graining make a difference to whether ‘Woods’ is a better guess than ‘Woods or Mickelson’? According to Dorst and Mandelkern’s account, the answer turns out to be ‘yes’. Their account predicts that as we increase the value of n —that is, as we increase the number of competitors—‘Woods or Mickelson’ eventually overtakes ‘Woods’ as the best guess.

To illustrate this, let us suppose that $J = 5$. Given this, we have seen that ‘Woods’ has a higher expected Dorst-Mandelkern answer-value than ‘Woods or Mickelson’ when $n = 1$ (this corresponds to the original scenario in which ‘Other’ is the only other option). But once we start increasing the value of n , ‘Woods or Mickelson’ eventually overtakes ‘Woods’ (as shown in Figure 3). The same thing can be shown to happen regardless of how high we choose the J -value to be. That is to say, no matter how much you value specificity, ‘Woods or Mickelson’ will eventually overtake ‘Woods’ once the number of competitors becomes sufficiently high.¹²

¹¹ Calculation: $E_{DM}(Woods) > E_{DM}(Woods\ or\ Mickelson) \iff$
 $.98 \cdot J^{(3-1)/3} > .99 \cdot J^{(3-2)/3} \iff$
 $J > 1.03.$

¹² Here is another way of putting essentially the same point: as we increase the number n , we can ask how much we need to increase the J -value to ensure that ‘Woods’ remains a better guess than ‘Woods or Mickelson’. The answer turns out to be that the minimum J -value required to ensure that $E_{DM}(Woods) > E_{DM}(Woods\ or\ Mickelson)$ approaches infinity as n approaches infinity. Here is the relevant calculation:

$$\begin{aligned}
 E_{DM}(Woods) > E_{DM}(Woods\ or\ Mickelson) & \iff \\
 .98 \cdot J^{(2+n+n-1)/(2+n)} > .99 \cdot J^{(2+n-2)/(2+n)} & \iff \\
 J > (.99/.98)^{2+n}, &
 \end{aligned}$$

which approaches infinity as n approaches infinity.

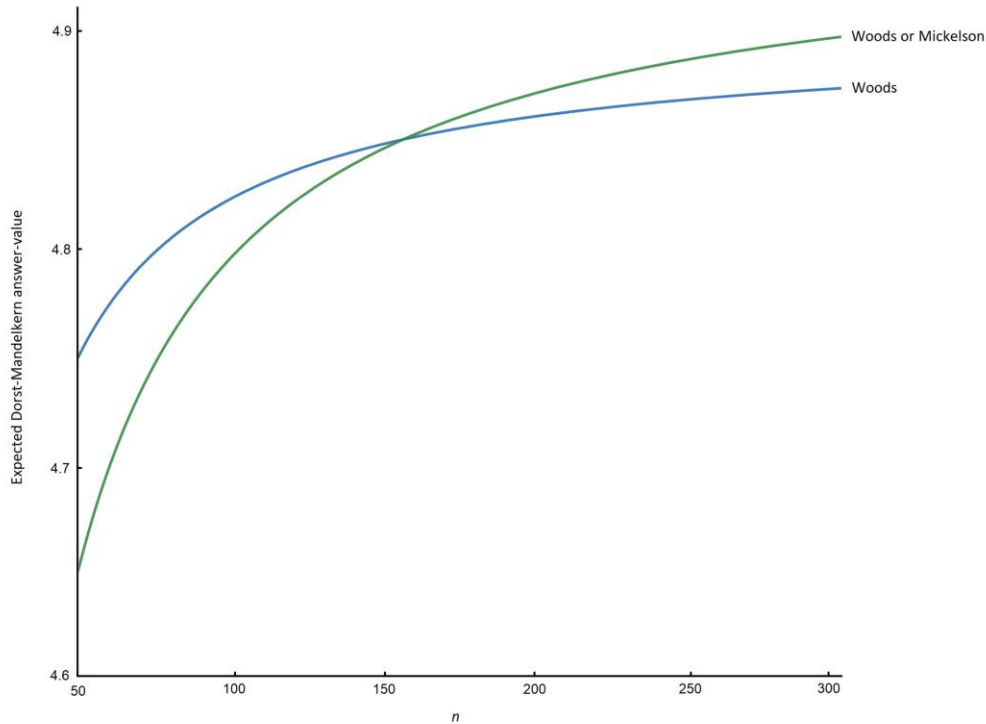


Figure 3: Expected Dorst-Mandelkern answer-value of ‘Woods’ and ‘Woods or Mickelson’, depending on n (setting $J = 5$ as an illustration).

Again, this strikes me as an odd result. We have just said that ‘Woods’ looks considerably more specific than ‘Woods or Mickelson’ while being only slightly less probable. So if you value specificity to a high enough degree, it would seem perfectly natural to guess ‘Woods’ rather than ‘Woods or Mickelson’, regardless of how many other players happen to share the residual 1% of the probability mass.

I want to suggest that this is an instance of a more general problem. The more general problem is that, on Dorst and Mandelkern’s account, whether one guess is better than another *depends on which other guesses are available*. In other words, Dorst and Mandelkern’s account fails to satisfy the following constraint:

Independence of Irrelevant Alternatives (for Guessing): If p is a better guess than q relative to Q , then for any question Q^* such that (i) the set of complete answers to Q which are compatible with p is identical to the set of complete answers to Q^* which are compatible with p , and (ii) the set of complete answers to Q which are compatible with q is identical to the set of complete answers to Q^* which are compatible with q , then p is a better guess than q relative to Q^* .

Put differently, whether one guess is better than another relative to a given accuracy-specificity tradeoff should only depend on the probabilities of the guesses and the number of complete answers with which they are compatible, not on the size of the question. This constraint mirrors a more familiar principle from rational choice theory—call it “IIA for

Preferences” to distinguish it from “IIA for Guessing”—which says that whether one option is preferable to another shouldn’t depend on which other options are available. For example, according to IIA for Preferences, whether you prefer Crème Brûlée to Tarte Tatin shouldn’t depend on whether there is also Mille-feuille on the dessert menu.¹³ Similarly, according to IIA for Guessing, whether one guess is better than another shouldn’t depend on which other guesses are available. For example, whether ‘Woods’ is a better guess than ‘Woods or Mickelson’ shouldn’t depend on how many other players happen to share the residual 1% of the probability mass.

Like many who have found IIA for Preferences to be an intuitive constraint on rational preferences, I find IIA for Guessing to be an intuitive constraint on good guesses. This is not to say that the two principles stand or fall together. Indeed, one of the most common challenges to IIA for Preferences does not, I think, have similar force against IIA for Guessing. Consider the following putative counterexample to IIA for Preferences, due to Sen (1993, p. 501):

Polite Cake-eating: You are choosing between slices of cake. You are trying to choose as large a slice as possible, subject to not choosing the very largest (you don’t want to appear greedy). Accordingly, when choosing from the menu {Small, Medium}, you prefer the small slice over the medium slice, but when choosing from the menu {Small, Medium, Large}, you prefer the medium slice over the small one.

As has often been pointed out, it is unclear whether a case like this really constitutes a genuine counterexample to IIA for Preferences, or whether it just shows that we need to be careful about how to individuate the available options.¹⁴ But however this may be, the parallel challenge to IIA for Guessing is much less compelling. Consider the following analogue of Sen’s example:

Cheap Guesses: You are guessing which horse is going to win the race. You are trying to choose a likely winner, subject to not picking the very biggest favorite (you consider that to be “too cheap”). Accordingly, when choosing from {Cody, Benji}, you prefer guessing ‘Cody’ over ‘Benji’, but when choosing from {Cody, Benji, Ajax}, you prefer guessing ‘Benji’ over ‘Cody’.

¹³ Here I’m deviating slightly from traditional formulations of IIA, which say that whether one option is preferred to another should not depend on the *removal* of irrelevant alternatives—see, e.g. Chernoff (1954), Sen (1969; 1971), and also Arrow (1951), though his principle of the same name concerns the relationship between individual and social preferences rather than the relationship amongst an individual’s preferences. On the present formulation, IIA also says that whether one option is preferred to another should not depend on the *addition* of irrelevant alternatives, thereby making the present formulation equivalent to the so-called “Weak Axiom of Revealed Preference” which is often taken as a consistency requirement on an individual’s preferences (cf. Samuelson 1938; 1948).

¹⁴ For example, rather than describing an option as “medium piece of cake,” one might describe it as “medium piece of cake, which is/is not the largest one available,” in which case the Polite Cake-Eating case becomes consistent with IIA for Preferences. See Rulli and Worsnip (2016) for a discussion of this line of response.

Although this case has the surface structure of a counterexample to IIA for Guessing, it seems clear that you are not really expressing your best guess here. Rather, you are performing a kind of speech act which may be influenced by various pragmatic factors having nothing to do with the accuracy or specificity of your guess (say, an aversion against making guesses that are hard to pronounce, or an aversion against making guesses that remind you of your ex-partner).¹⁵

There is, however, a different way one might try to push back against IIA for Guessing. Consider the following variation on the golf case:

Woods	Mickelson	P_1	P_2	...	P_n
50%	40%	10%/n	10%/n		10%/n

Again, we want to compare ‘Woods’ and ‘Woods or Mickelson’, for different values of n . For concreteness, let us focus on $n = 1$ vs. $n = 100$:

$n = 1$	Woods	Mickelson	P_1		
	50%	40%	10%		
$n = 100$	Woods	Mickelson	P_1	P_2	...
	50%	40%	0.1%	0.1%	...

Suppose you consider ‘Woods’ to be a better guess than ‘Woods or Mickelson’ when $n = 1$. We can then ask: should you also consider ‘Woods’ to be a better guess than ‘Woods or Mickelson’ when $n = 100$? According to IIA for Guessing, the answer is ‘yes’ (assuming that your preferred accuracy-specificity tradeoff doesn’t change). But you might think that it can be reasonable to guess ‘Woods or Mickelson’ when $n = 100$, even if you prefer to guess ‘Woods’ when $n = 1$. After all, when $n = 100$, Tiger Woods and Phil Mickelson “stand out” as the two clear favorites in a way that they don’t when $n = 1$. In other words, there is a sense in which ‘Woods and Mickelson’ becomes an increasingly salient guess as n increases. Given this, it might seem natural to switch guesses from ‘Woods’ to ‘Woods or Mickelson’ as n increases from 1 to 100.

Although this way of thinking has some intuitive appeal, I find it difficult to uphold the intuition on reflection. Imagine that you and your friend are about to watch a golf tournament featuring three professional players: Woods (50%), Mickelson (40%), and McIlroy (10%). During the warm-up session, your friend asks: “Who do you think is going to win?” You answer: “Woods.” A moment later, a sad announcement is made: McIlroy is forced to withdraw due to a knee injury. Instead, the organizers have decided to let 100 amateur golfers compete, each with a 0.1% chance of winning. In light of these changes,

¹⁵ Here I’m in broad agreement with Dorst and Mandelkern who distinguish the *speech act* of guessing from the *cognitive attitude* of guessing, only the former of which may be influenced by pragmatic factors.

your friend asks again: “Who do you *now* think is going to win?” You answer: “Woods or Mickelson.”

This strikes me as an odd response; and the reason why it strikes me as an odd response, I think, is that it’s hard for me to see how it could matter whether the residual 10% of the probability mass is distributed across a hundred players or just one, even if this makes a difference to how clearly Woods and Mickelson stand out as the two biggest favorites. Perhaps not everyone will share this intuition. I certainly don’t purport to have given a decisive argument for IIA for Guessing here. But I think enough can be said in its favor to make it worthwhile exploring whether we can formulate an account of good guesses that satisfies it.

4. An Alternative Proposal

I now want to offer an alternative account of good guesses, which avoids the concerns raised. The account is in many ways congenial to Dorst and Mandelkern’s account: it also says that good guesses optimize a tradeoff between accuracy and specificity. But the underlying measure of specificity is different. Rather than measuring specificity in terms of the proportion of complete answers a guess rules out, I propose to measure specificity in terms of the following log-ratio:

Specificity: The specificity of a guess, p , in response to a question, Q , is given by:

$$S(p) = \log \frac{|Q|}{c_p},$$

where c_p is the number of complete guesses which are compatible with p :
 $c_p = |\{q \in Q: p \cap q \neq \emptyset\}|$.

Before we explore the consequences of adopting this measure of specificity, I want to highlight some of its properties.

The first thing to observe is that $S(p)$ is a decreasing function of c_p , and an increasing function of $|Q|$: when $c_p = |Q|$, $S(p) = 0$, and when $c_p = 1$, $S(p) = \log(|Q|)$, which is a quantity that approaches infinity as $|Q|$ approaches infinity. In this respect, our new specificity measure is similar to that of Dorst and Mandelkern: both measures faithfully capture the idea that a specific guess is one that narrows down the space of possibilities under consideration, and that how far a guess can narrow down the space of possibilities depends on the size of the question under consideration.

The second thing to observe is that our new specificity measure bears resemblance to Shannon’s (1948) classic measure of information, according to which the amount of information contained in a proposition, p , is given by $\log(1/P(p))$. When the probability distribution over the complete answers is flat—that is, when each complete answer has a

probability of $1/|Q|$ —the probability of p is given by $c_p/|Q|$, which means that the specificity of p reduces to $S(p) = \log(1/P(p))$. So when the probability distribution is flat, the specificity of a proposition is identical to its Shannon information.¹⁶

This also means that our new measure of specificity inherits certain features of Shannon’s measure of information. For example, $S(p)$ has no finite upper bound: it ranges from 0 to $\log(|Q|)$, which is a quantity that approaches infinity as $|Q|$ approaches infinity. By contrast, Dorst and Mandelkern’s measure of specificity does have a finite upper bound: it ranges from 0 to $(|Q| - 1)/|Q|$, which is a quantity that approaches 1 as $|Q|$ approaches infinity. Although I won’t place much weight on this difference, it seems to me that, insofar as we have an intuitive grip on the notion of specificity, there should be no upper limit on how specific a guess can in principle be, just as there is no upper limit on how informative a proposition can in principle be. But as I said, I won’t rest anything much on this point. The proposed measure of specificity will earn its keep by helping us solve the problems raised in the previous section.

With our new specificity measure in hand, we can define the answer-value of a guess as follows:

Answer-value: The *answer-value* of a guess, p , in response to a question, Q , is given by:

$$V^+(p) = J^{S(p)} = J^{\log(|Q|/c_p)},$$

where $J \geq 1$.

Accordingly, the *expected* answer-value of p becomes:

$$E(p) = P(p) \cdot J^{S(p)} = P(p) \cdot J^{\log(|Q|/c_p)}$$

This is the quantity that, I propose, one should try to maximize when forming one’s guess: a guess, p , is permissible iff there is no alternative guess, q , such that $E(q) > E(p)$.

How does this account compare to that of Dorst and Mandelkern? Both accounts satisfy No Accuracy-Dominance and No Specificity-Dominance, since both accounts are based on a truth-directed and specificity-directed measure of answer-value. But unlike Dorst and Mandelkern’s account, our new account satisfies Neutrality and IIA for Guessing as well, whereas it doesn’t satisfy Optionality. Let us consider each of these constraints in turn.

¹⁶ Someone might wonder why we can’t simply identify the specificity of p with its Shannon information. The reason for this is that the specificity of p —unlike the informativeness of p —is supposed to be independent of how probable p is. For example, the guesses ‘Ajax’ and ‘Ember’ should count as being equally specific, even though ‘Ajax’ is much more probable than ‘Ember’. By contrast, the notion of informativeness which Shannon sought to capture is supposed to be inversely related to a proposition’s probability, the idea being that it isn’t very surprising to learn that p is true if you’re already quite confident of p . So, for example, ‘Ajax’ should count as being much less informative than ‘Ember’, even though both propositions are equally specific.

The following result shows that our new account satisfies Neutrality (see Appendix A for a proof):

Point of Neutrality: Let p and q be arbitrary guesses in response to a question, Q , and suppose that the probability distribution over the complete answers to Q is flat. Then $E(p) = E(q)$ iff $J = e$.

We can think of $J = e$ as the “point of neutrality” from which one can become more risk-averse or risk-seeking by lowering or raising the J -value.¹⁷ For example, in the die roll case, the best guess will be the entire disjunction ‘1 v 2 v 3 v 4 v 5 v 6’ when $J < e$; the six complete answers (‘1’, ‘2’, ‘3’, etc.) will tie for best when if $J > e$; and all guesses will be equally good when $J = e$ (see Figure 4 for an illustration).

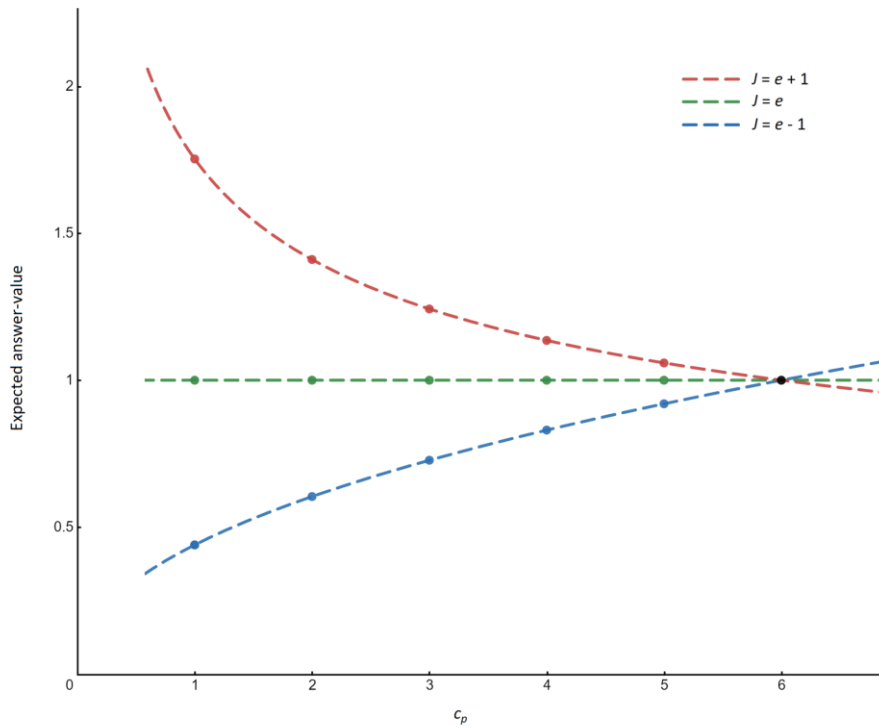


Figure 4: Expected answer-value of various die roll guesses, depending on J .

Next, consider IIA for Guessing. We have seen that Dorst and Mandelkern’s account violates IIA for Guessing in cases like the following:

Woods	Mickelson	P_1	P_1	...	P_n
98%	1%	$1\%/n$	$1\%/n$		$1\%/n$

¹⁷ There is nothing special about Euler’s constant here, except that it is the chosen log base. If we had chosen a different log base, say 10, the point of neutrality would instead have been $J = 10$.

Even if ‘Woods’ starts out as a better guess than ‘Woods or Mickelson’ for low values of n , Dorst and Mandelkern’s account implies that, as n increases, ‘Woods or Mickelson’ eventually overtakes ‘Woods’. By contrast, our new account implies that if ‘Woods’ starts out as a better guess than ‘Woods or Mickelson’ for low values of n , it remains better no matter how much we increase the value of n (as illustrated by Figure 5). This can be shown to hold in full generality: whether one guess has a higher expected answer-value than another depends only on the probabilities of the guesses and the number of complete answers with which they are compatible, as required by IIA for Guessing (see Appendix B for a proof).

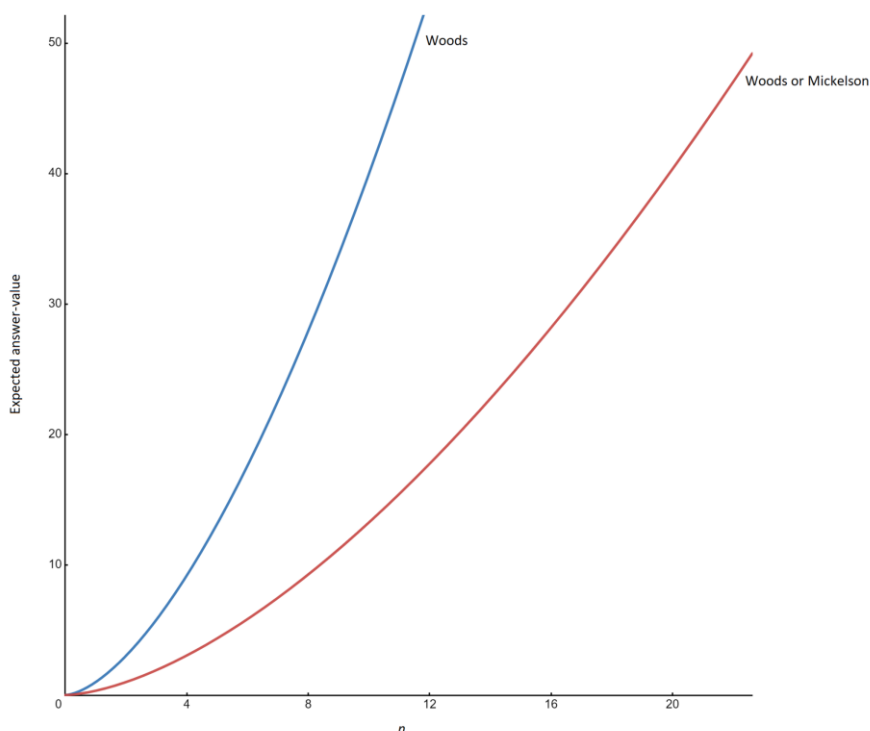


Figure 5: Expected answer-value of ‘Woods’ and ‘Woods or Mickelson’, depending on n (setting $J = 5$ as an illustration).

Finally, consider Optionality. Like Dorst and Mandelkern’s account, our new account predicts a fairly high degree of optionality when it comes to deciding how many complete answers to include in one’s guess. Here is a simple example:

Ajax	Benji	Cody
50%	30%	20%

On our new account, each of the three filtered guesses (‘Ajax’, ‘Ajax or Benji’, and ‘Ajax, Benji or Cody’) can be made to have the highest expected answer-value by choosing the right J -value: ‘Ajax’ is best when $J < 1.73$; ‘Ajax, Benji or Cody’ is best when $J > 1.97$; and ‘Ajax or Benji’ is best otherwise (see Figure 6).

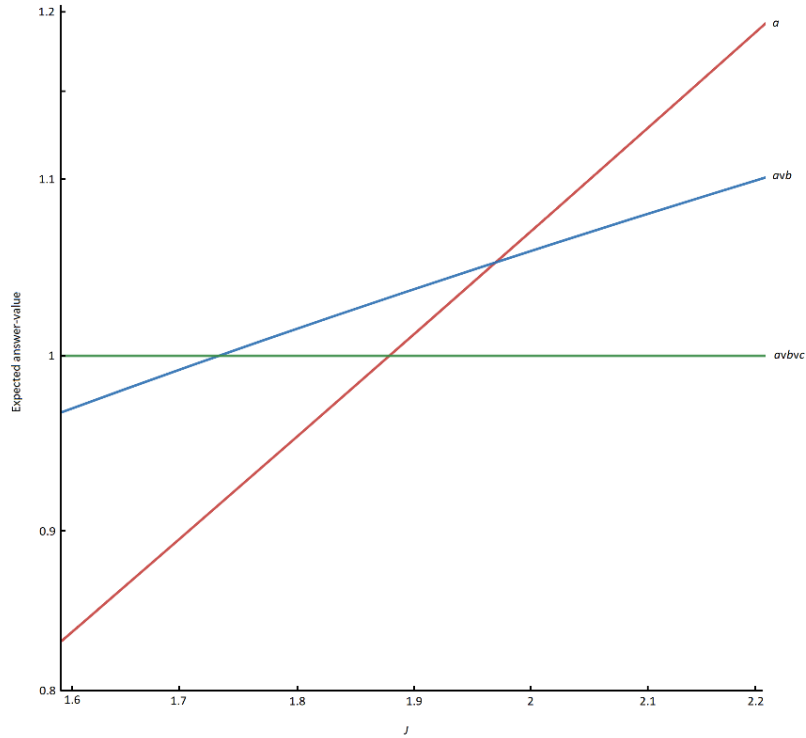


Figure 6: Expected answer-value of ‘Ajax’, ‘Ajax or Benji’, and ‘Ajax, Benji or Cody’, depending on J , where the probability distribution is {Ajax: 50%, Benji: 30%, Cody: 20%}.

However, unlike Dorst and Mandelkern’s account, our new account doesn’t satisfy Optionality in full generality. Consider what happens if we change the probability distribution in such a way that Benji and Cody have almost the same winning chances:

Ajax	Benji	Cody
50%	26%	24%

On Dorst and Mandelkern’s account, it will still be possible to pick a J -value relative to which ‘Ajax or Benji’ is a permissible guess.¹⁸ But no such J -value exists on our new account.¹⁹ Why

¹⁸ As is easily verified, the following inequalities are jointly satisfied iff $2.28 \leq J \leq 3.51$:

$$\begin{aligned}
 E_{\text{DM}}(\text{Ajax}) \leq E_{\text{DM}}(\text{Ajax or Benji}) & \quad \& \quad E_{\text{DM}}(\text{Ajax, Benji or Cody}) \leq E_{\text{DM}}(\text{Ajax or Benji}) \Leftrightarrow \\
 .50 \cdot J^{(3-1)/3} \leq (.5 + .26) \cdot J^{(3-2)/3} & \quad \& \quad 1 \leq (.5 + .26) \cdot J^{(3-2)/3}
 \end{aligned}$$

¹⁹ If ‘Ajax or Benji’ were a permissible guess, the following two inequalities would have to hold:

$$\begin{aligned}
 E(\text{Ajax}) \leq E(\text{Ajax or Benji}) & \quad \& \quad E(\text{Ajax, Benji or Cody}) \leq E(\text{Ajax or Benji}) \Leftrightarrow \\
 .5 \cdot J^{\log(3/1)} \leq .77 \cdot J^{\log(3/2)} & \quad \& \quad 1 \cdot J^{\log(3/3)} \leq .77 \cdot J^{\log(3/2)}
 \end{aligned}$$

Yet it is simple to show that no J -value renders both inequalities true. So, on our new account, no accuracy-specificity tradeoff makes it permissible to include exactly two complete answers in your guess.

not? The mechanism is illustrated in Figure 7: for low J -values, ‘Ajax, Benji or Cody’ is a better guess than both ‘Ajax’ and ‘Ajax or Benji’ in virtue of having a higher probability. But once we start increasing the J -value, ‘Ajax’ and ‘Ajax or Benji’ both eventually overtake ‘Ajax, Benji or Cody’ in virtue of being more specific. And, as it turns out, ‘Ajax’ closes the gap to ‘Ajax, Benji or Cody’ more quickly than does ‘Ajax or Benji’. As a result, ‘Ajax or Benji’ never gets to be the best guess: it always loses to either ‘Ajax’ or ‘Ajax or Benji’ (or both).

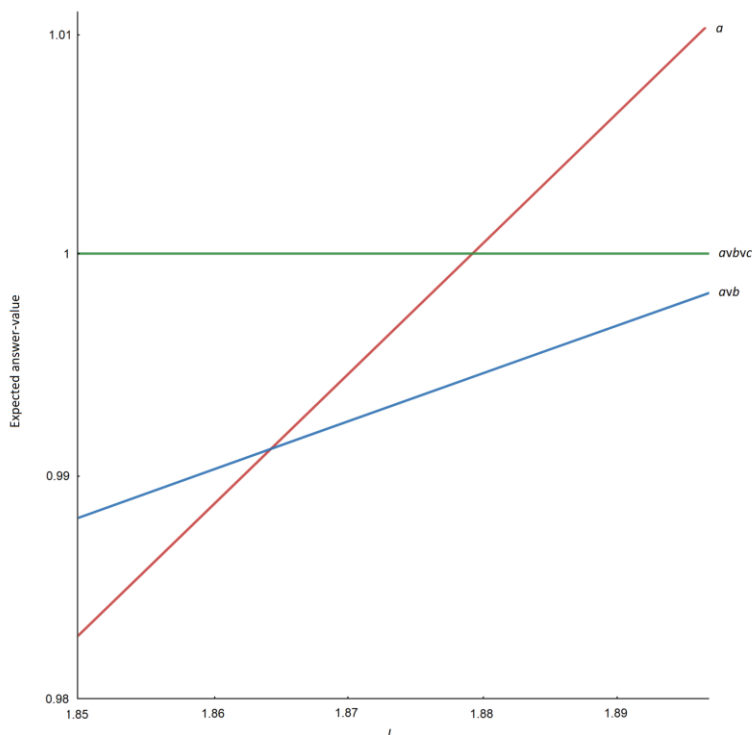


Figure 7: Expected answer-value of ‘Ajax’, ‘Ajax or Benji’, and ‘Ajax, Benji or Cody’, depending on J , where the probability distribution is {Ajax: 50%, Benji: 26%, Cody: 24%}.

Are such violations of Optionality a feature or a bug of the proposed account? Like Dorst and Mandelkern, I find it very plausible that an account of good guesses should be quite permissive when it comes to deciding how many complete answers to include in one’s guess. But I doubt that something as strong as Optionality holds in full generality. Consider what happens if we further change the horse race example so that Benji and Cody end up with exactly the same winning chances:

Ajax	Benji	Cody
50%	25%	25%

As Dorst and Mandelkern themselves observe, there is something rather odd about guessing ‘Ajax or Benji’ in such a case. What explains this? Dorst and Mandelkern’s suggestion is this: although it is *permissible* to guess ‘Ajax or Benji’, people tend to avoid making guesses that

crosscut “clusters” of complete answers with similar probabilities. So, for example, people tend to prefer J -values that allow them to guess either ‘Ajax’ or ‘Ajax, Benji or Cody’ in a case like that above.

This is perhaps a natural enough explanation of the intuitive oddness of guessing ‘Ajax or Benji’. But another natural explanation, it seems to me, is that guessing ‘Ajax or Benji’ is indeed prohibited, as predicted by our new account. I’m not sure how to adjudicate between these rival explanations in isolation from other supporting considerations, such as those laid out previously in the paper. But it seems to me that the intuitive case for Optionality is much less compelling than one might have initially thought.

There is, however, a residual worry that one might have about the violations of Optionality engendered by our new account. Consider a schematic version of the horse race example just discussed (where $.25 < x < .5$):

a	b	c
50%	x	$50\% - x$

As we have seen, our new account violates Optionality when $x = 26\%$, but not when $x = 30\%$. Given this, we should expect that there exists a threshold, t , such that Optionality is violated iff $x < t$. This is precisely what we find: in the case at hand, the threshold turns out to be $t \approx .274$.²⁰ But what could explain the existence of such a seemingly arbitrary cut-off? That is the worry.

I think the worry can be at least partly assuaged by observing that the threshold in question is not introduced by hand, but rather falls out as a consequence of an independently motivated account of good guesses. The mechanism by which the threshold arises is illustrated in Figure 8: when the probabilities of b and c are sufficiently far apart, $a \vee b$ overtakes $a \vee b \vee c$ before a does. But as the probabilities of b and c get closer and closer to each other, a gets closer and closer to overtaking $a \vee b \vee c$ before $a \vee b$ does. And eventually, when the probabilities of b and c get sufficiently close, a does indeed overtake $a \vee b \vee c$ before $a \vee b$ does, with the result that Optionality is violated.

Thus, there is a sense in which the threshold is not arbitrary at all: it is fully explicable by how the proposed account works. Of course, this won’t bring peace of mind to those who were hoping for a purely pre-theoretic explanation of why the threshold in question exists. But I doubt that pre-theoretic considerations will take us very far in deciding whether a threshold of this nature is acceptable. Either way, I take comfort in the fact that the threshold is not a result of gerrymandering, but a consequence of an independently motivated account of good guesses.

²⁰ As is easily verified, the following inequalities can be jointly satisfied by some value of J iff $.25 \leq x \leq .274$:

$$\begin{array}{llll}
 E(a) \leq E(a \vee b) & \& E(a \vee b \vee c) \leq E(a \vee b) & \Leftrightarrow \\
 .50 \cdot J^{(3-1)/3} \leq (.5 + x) \cdot J^{(3-2)/3} & \& 1 \leq (.5 + x) \cdot J^{(3-2)/3} &
 \end{array}$$

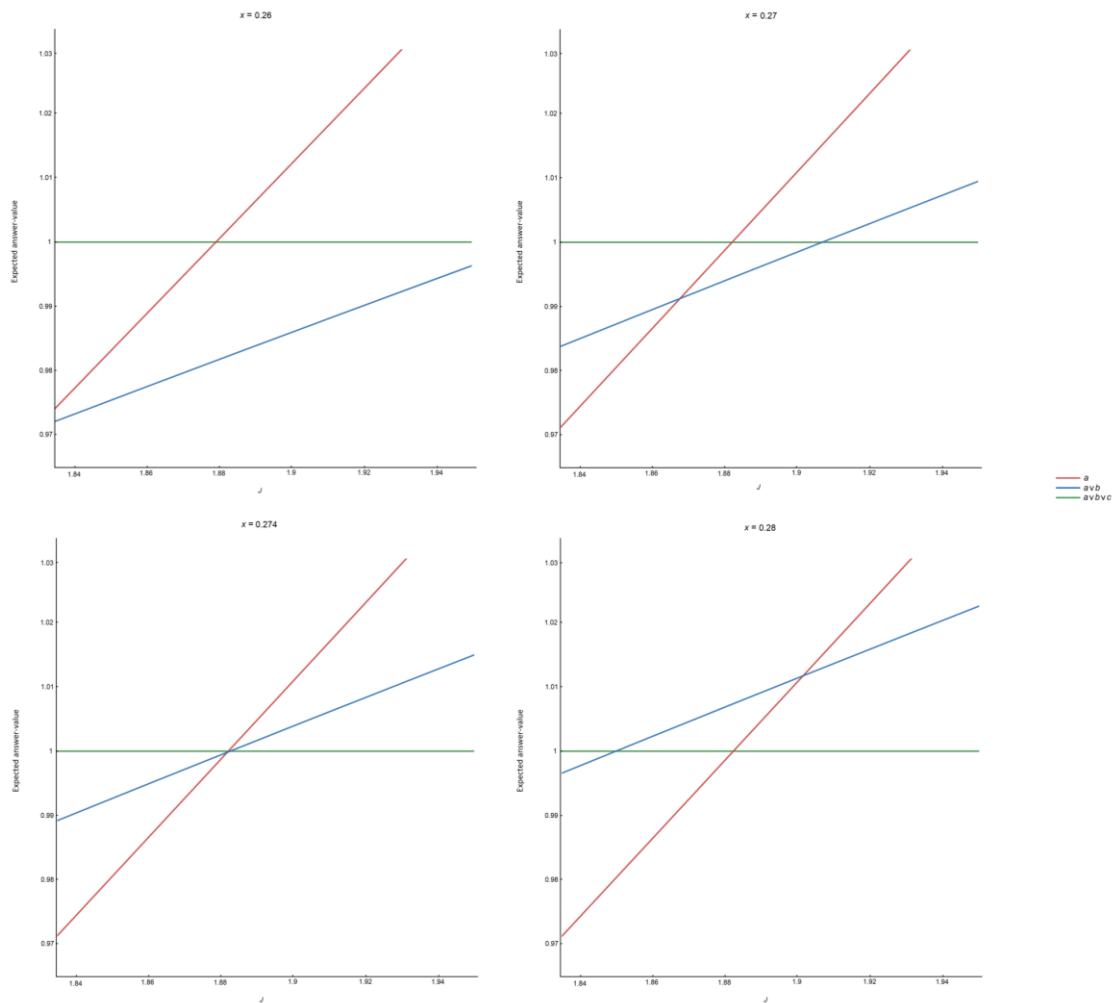


Figure 8: Expected answer-value of a , avb , and $avbvc$, depending on J , for different values of x .

In sum, although I certainly haven't argued decisively against Optionality here, I am inclined to think that we should treat the putative counterexamples to Optionality as an interesting, and perhaps somewhat surprising, consequence of an account of good guesses whose main attractions are that it gives us a simple and elegant way to satisfy Neutrality and IIA for Guessing.

5. Conclusion

There is much to like about Dorst and Mandelkern's account of good guesses. It is based on a simple and attractive idea, that good guesses optimize a tradeoff between accuracy and specificity, and it offers a precise implementation of this idea which lends itself to systematic investigation. Nonetheless, I have argued that their implementation fails to satisfy some plausible constraints on good guesses, and I have offered an alternative implementation that satisfies these constraints. Although the positive proposal of the paper differs from Dorst

and Mandelkern's proposal in some important respects, it retains the basic idea that good guesses optimize a tradeoff between accuracy and specificity. Thus, I would like to think that the considerations put forth in this paper do not undermine Dorst and Mandelkern's overall project, but in fact vindicate it.

Appendices

A. Proof of Point of Neutrality

Let Q be a question with n complete answers, and suppose the probability distribution over the complete answers is flat:

p_1	p_2	p_3	...	p_n
$1/n$	$1/n$	$1/n$		$1/n$

We must show that, for any p and q , $E(p) = E(q)$ iff $J = e$. We begin by taking the derivative of $E(p_1 \vee p_2 \vee \dots \vee p_i)$ with respect to i :

$$E(p_1 \vee p_2 \vee \dots \vee p_i) = \frac{i}{n} \cdot J^{\log\left(\frac{n}{i}\right)} \Rightarrow$$

$$\frac{dE(p_1 \vee p_2 \vee \dots \vee p_i)}{di} = - \frac{(\log(J) - 1) \cdot J^{\log\left(\frac{n}{i}\right)}}{n}$$

We then solve the equation $\frac{dE(p_1 \vee p_2 \vee \dots \vee p_i)}{di} = 0$ for J :

$$- \frac{(\log(J) - 1) \cdot J^{\log\left(\frac{n}{i}\right)}}{n} = 0 \Leftrightarrow$$

$$\log(J) - 1 = 0 \Leftrightarrow$$

$$J = e.$$

What this shows is that $E(p_1 \vee p_2 \vee \dots \vee p_i)$ is a constant function of i iff $J = e$. Thus, $E(p) = E(q)$ iff $J = e$.

B. Proof of IIA for Guessing

Let p and q be arbitrary answers to a question of size n , and let $d(n)$ be the difference between the expected answer-values of p and q :

$$d(n) := E(p) - E(q) = P(p) \cdot J^{\log\left(\frac{n}{|p|}\right)} - P(q) \cdot J^{\log\left(\frac{n}{|q|}\right)}.$$

To show that the proposed account satisfies IIA for Guessing, it suffices to establish the following:

Case 1: If $d(2) > 0$, then $d(n) = 0$ has no solutions for $n \geq 2$.

Case 2: If $d(2) = 0$, then $d(n) = 0$, for all $n \geq 2$.

Case 1: Suppose $d(2) > 0$. Then we have:

$$P(p) \cdot J^{\log\left(\frac{2}{|p|}\right)} - P(q) \cdot J^{\log\left(\frac{2}{|q|}\right)} > 0 \Leftrightarrow$$

$$\frac{P(p)}{\log(|p|)} > \frac{P(q)}{\log(|q|)}.$$

We then try to solve the equation $d(n) = 0$:

$$P(p) \cdot J^{\log\left(\frac{n}{|p|}\right)} - P(q) \cdot J^{\log\left(\frac{n}{|q|}\right)} = 0 \Leftrightarrow$$

$$\frac{P(p)}{\log(|p|)} = \frac{P(q)}{\log(|q|)},$$

which is inconsistent with the inequality above. Thus, if $d(2) > 0$, then $d(n) = 0$ has no solutions for $n \geq 2$.

Case 2: Suppose $d(2) = 0$. Then we have:

$$P(p) \cdot J^{\log\left(\frac{2}{|p|}\right)} - P(q) \cdot J^{\log\left(\frac{2}{|q|}\right)} = 0 \Leftrightarrow$$

$$\frac{P(p)}{\log(|p|)} = \frac{P(q)}{\log(|q|)}.$$

As we have seen, this equality is equivalent to the claim that $d(n) = 0$, for all $n \geq 2$.

Acknowledgments: For helpful feedback, I am grateful to the philosophy reading group at National University of Singapore, Kevin Dorst, Matthew Mandelkern, two anonymous referees for *Philosophical Studies*, and especially Zach Barnett. This research was supported by a Tier 1 research grant from the Ministry of Education in Singapore.

References

- Arrow, K. (1951 [1963]): *Social Choice and Individual Values*. New York: Wiley. 2nd ed., 1963.
- Chernoff, H. (1954): “Rational Selection of Decision Functions.” *Econometrica* 22, pp. 422–43.
- Dorst, K. and M. Mandelkern (forthcoming): “Good Guesses.” *Philosophy and Phenomenological Research*. Early view.
- Groenendijk, J. and Stokhof, M. (1984): *Studies on the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, University of Amsterdam.

- Hamblin, C. L. (1973): “Questions in Montague English.” *Foundations of Language* 10, pp. 41–53.
- Hawthorne, J., D. Rothschild, and L. Spectre (2016): “Belief is Weak.” *Philosophical Studies* 173, pp. 1393–1404.
- Holguín, B. (2022): “Thinking, Guessing, and Believing.” *Philosophers’ Imprint* 22(6).
- James, W. (1896 [1956]): “The Will to Believe.” In: *The Will to believe and other essays in popular philosophy*. New York: Dover Publications, pp. 1–31.
- Rulli, T. and A. Worsnip (2016): “IIA, Rationality, and the Individuation of Options.” *Philosophical Studies* 173, pp. 205–21.
- Samuelson, P. (1938): “A Note on the Pure Theory of Consumer’s Behavior.” *Economica* 5, pp. 61–71.
- Samuelson, P. (1948): “Consumption Theory in Terms of Revealed Preference.” *Economica* 15, pp. 243–53.
- Sen, A. (1969): “Quasi-Transitivity, Rational Choice, and Collective Decisions.” *The Review of Economic Studies* 36, pp. 381–93.
- Sen, A. (1971): “Choice Functions and Revealed Preference.” *The Review of Economic Studies* 38, pp. 307–17.
- Sen, A. (1993): “Internal Consistency of Choice.” *Econometrica* 61, pp. 495–521.
- Shannon, C. (1948): “The Mathematical Theory of Communication.” *Bell System Technical Journal* 27, pp. 379–423.