Perceptuo-motor, cognitive, and description-based decision-making seem equally good

Andreas Jarvstad[a,b,1], Ulrike Hahn[a,c], Simon K. Rushton[a], and Paul A. Warren[d]

[a]School of Psychology, Cardiff University, Cardiff CF10 3AT, United Kingdom; [b]School of Experimental Psychology, University of Bristol, Clifton, Bristol BS8 1TU, United Kingdom; [c]Department of Psychological Sciences, Birkbeck, University of London, London WC1E 7HX, United Kingdom; and [d]School of Psychological Sciences, University of Manchester, Manchester M13 9PT, United Kingdom

Edited by Randy Gallistel, Rutgers University, Piscataway, NJ, and approved August 13, 2013 (received for review January 9, 2013)

Classical studies suggest that high-level cognitive decisions (e.g., choosing between financial options) are suboptimal. In contrast, low-level decisions (e.g., choosing where to put your feet on a rocky ridge) appear near-optimal: the perception–cognition gap. Moreover, in classical tasks, people appear to put too much weight on unlikely events. In contrast, when people can learn through experience, they appear to put too little weight on unlikely events: the description–experience gap. We eliminated confounding factors and, contrary to what is commonly believed, found results suggesting that (*i*) the perception–cognition gap is illusory and due to differences in the way performance is assessed; (*ii*) the description–experience gap arises from the assumption that objective probabilities match subjective ones; (*iii*) people's ability to make decisions is better than the classical literature suggests; and (*iv*) differences between decision-makers are more important for predicting peoples' choices than differences between choice tasks.

Decades of research on the human ability to make choices suggests that humans are suboptimal decision-makers (1–5). In classical studies demonstrating suboptimal choices, a description-based paradigm is typically used. You might be asked if you prefer option A "£4000 with a probability of 0.2," or option B "£3000 with a probability of 0.25" (3). Thus, you are asked to make choices on the basis of numerically described probability and value information. By asking people such questions, researchers have uncovered many departures from optimal decision-making. More recently, however, two other decision-making paradigms have produced results that seem to diverge dramatically from the classical ones.

First, in sharp contrast to the classical studies, recent studies of decision-making in perceptuo-motor (6, 7) and perceptual (8–11) domains typically report optimal or near-optimal decisions. Like the classical ones, these provide participants with numerical value information. However, participants do not receive explicit probability information as in the description-based paradigm. Instead, participants have to use internal, implicit estimates of probabilities (see below for an example). Thus, low-level decisions based on internal estimates of probability are near optimal, whereas high-level decisions based on verbally described numerical probabilities are suboptimal: the perception–cognition gap (12, 13).

Second, within cognitive contexts, "decisions from experience" dissociate from the classical "decisions from description": the description–experience gap (14, 15). In decisions from experience, participants typically sample two computer buttons, each generating monetary outcomes with some probability. One button might return £32 with a probability of 0.1 and the other returns £3 guaranteed (16). After sampling both buttons, participants choose which button to play for real money.* In this example, people are more likely to choose the certain option compared with when options are described numerically (as in the classical studies). In other words, people act as if they attach too much importance to unlikely events (overweighting) in classical tasks, whereas acting as if they attach too little importance to unlikely events (underweighting) in experience tasks.†

The underlying causes of both of these gaps have attracted considerable attention (14, 15, 18), but are far from resolved. As pointed out previously (18), the paradigms differ along a number of dimensions; any one of these could potentially explain away the gaps. Some of the differences were listed above, but there are many more. Classical experiments, for example, present participants with a select set of choice options whose expected outcomes typically differ very little. The options are selected such that particular choice patterns across pairs of decision-making problems are indicative of irrational choice. Normally, participants are not given feedback and make one-shot decisions (3). Some studies in the decisions-from-experience literature, in contrast, give participants trial-by-trial feedback, although some do not (14), and participants in perceptuo-motor studies typically make repeated choices with feedback (13).

Are the two gaps, or dissociations, a result of the human mind operating in fundamentally different ways across domains? Or do they result from methodological differences across experimental paradigms? For example, are perceptuo-motor decisions better than cognitive decisions because feedback is provided or because they are perceptuo-motor and not cognitive? One can approach such questions in at least two ways. One could study the effects of a factorial combination of all methodological differences across

Significance

Human decision-making seems fundamentally domain dependent. Sensory-motor decisions (e.g., where to put your feet on a rocky ridge) seem near-optimal, whereas decisions based on numerical information (e.g., choosing between financial options) seem suboptimal. Additionally, when people rely on information gained through experience, they make choices that are often the opposite of those they make when relying on described information. However, comparing results across domains on the basis of past results is difficult, because decision-making is studied very differently in different domains. We compared decision-making performance across domains under precisely matched conditions, finding evidence against the idea that fundamental dissociations exist. In fact, peoples' ability to make decisions seem rather good, although not perfect, in both sensory-motor and cognitive domains.

Author contributions: A.J., U.H., S.K.R., and P.A.W. designed research; A.J. performed research; A.J. and U.H. contributed new analytic tools; A.J. analyzed data; and A.J., U.H., S.K.R., and P.A.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: a.jarvstad@bristol.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1300239110/-/DCSupplemental.

*Variants of this sampling paradigm in which each sample is played out for real with varying levels of feedback are also used (14). Note also that ref. 16 paid participants in euros and not pound sterling.

†The decision from experience literature is very rich, and this feature is by no means the only phenomenon discovered and explored theoretically. For example, the literature consistently observes what appears to be dramatic undersampling. When allowed to sample freely before committing to one option, people sample very few times. Intuitively, this is a suboptimal strategy because it gives rise to poor estimates of the value of the underlying choice options. Recently, however, it was proposed that such undersampling may actually be beneficial as it amplifies the difference between choice options making discriminations between choice options easier (17).

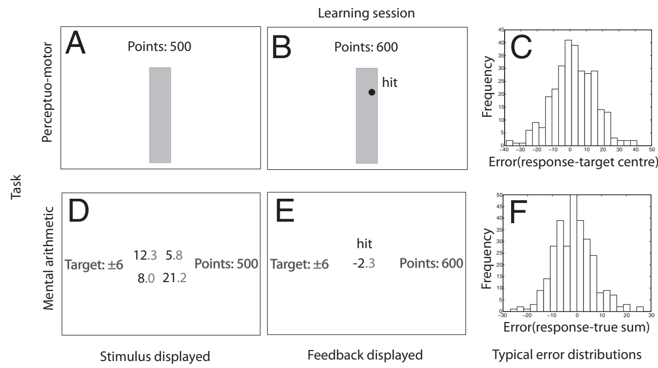

Fig. 1. Learning session. In both tasks, target width was not fixed but varied from trial to trial. (*A*) Pointing stimuli. The bar represents the target and "Points 500" illustrates the cumulative score information. (*B*) Pointing feedback included explicit hit/miss information as well as a high-contrast marker indicating where the screen was hit (black disk here). (*C*) A pointing error distribution (participant 10). (*D*) Mental arithmetic stimuli. The task involved summing the four central numbers. The judged sum had to be within the limits of the target (±6 here) to be scored a hit. The total number of points earned was displayed to the right. (*E*) Mental arithmetic feedback. Feedback included the error (difference between the true and the judged sum, −2.3 here) and explicit hit/miss information. (*F*) An arithmetic error distribution (participant 10).

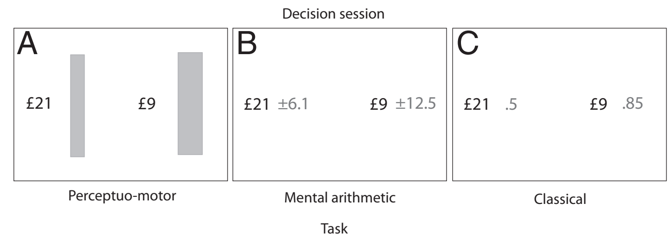

Fig. 2. Decision session. *A*–*C* illustrate the three decision tasks. (*A*) Two choice options with probabilities replaced by pointing targets. (*B*) Two choice options with probabilities replaced by arithmetic targets. (*C*) Two classical choice options with numerical probabilities. Participants choose between the options by pressing the left or the right button on a computer mouse.

the three paradigms. Alternatively, one could try to eliminate any differences not due to the dimension of interest. To the extent that differences across perception and cognition, or description and experience, persist when tasks have been matched, they must reflect fundamental differences across domains. Here we took the latter approach.

To explore whether there really is robust evidence for gaps, we compared three types of decision under precisely matched conditions. The first decision type was classical, with explicit, numerical descriptions of probability information. In the second and third types, explicit probability information was replaced by implicit estimates arising from experience with either a low-level perceptuo-motor task (a standard pointing task) or with a high-level cognitive task (a mental arithmetic task).[‡]

Learning and Decision Phase

The present study made use of the idea that classical decision tasks can be altered to require participants to use their own internal estimates of probabilities (19, 20) and was inspired by experiment 2 in ref. 20. Participants first learned about their ability to perform a standard low-level perceptuo-motor (20) task and a novel high-level mental arithmetic task. They later had to use the acquired knowledge to make decisions.

The perceptuo-motor learning task involved pointing toward targets displayed on a computer screen. Targets varied in size from trial to trial. Participants' hands were positioned in front of the screen before each trial. Participants had to hit the displayed target (illustrated by the shaded bar in Fig. 1*A*). Target hits were rewarded with 100 points, misses were not penalized, and late responses were penalized with −700 points. Feedback consisted of a high-contrast disk showing where the screen was hit (Fig. 1*B*).

In the arithmetic learning task, participants had to sum four numbers (randomly drawn on each trial, illustrated by the central numbers in Fig. 1*D*). The goal was to estimate the sum of the four numbers within a target tolerance (which varied across trials). In the example in Fig. 1 *D* and *E*, the true sum was 47.3 and the participant guessed 45. The difference between the response and the true sum is −2.3 (Fig. 1*E*). Because this difference was within the target tolerance (±6), a hit was scored. Points were rewarded as in the perceptuo-motor task. Participants responded with an on-screen keypad (not shown here).

In the learning phase, participants were instructed to earn as many points as possible and were told that learning about their own performance was at least as important as earning points. It was emphasized that improved knowledge would enable them to make better decisions in the decision-making phase and that those decisions would have real financial implications.

By collating participants' responses over many pointing and mental arithmetic trials, we can generate error distributions (Fig. 1 *C* and *F*), describing the accuracy and precision of individual participants' responses. These distributions can be used to predict participants' chances of hitting targets of varying widths.[§]

In the decision-making phase, participants had to choose between pairs of options (Fig. 2 *A*–*C*). Each choice option was composed of reward and probability information. The probability information was in one of three formats: low level (A), high level (B), and classical (C). In the former two cases, the width of the target determined the probability of obtaining the reward. The narrower the pointing target (A) and the narrower the arithmetic target (B), the lower the probability of obtaining the reward. The probability of obtaining the reward in the classical format (C) was described numerically. In all three tasks, participants simply indicated which option they preferred.

The rewards and the likelihoods of obtaining them were matched across the three tasks. Options for the classical task were created by drawing randomly from a range of probabilities and rewards (see *Methods* for details). Using the empirical error distributions from the learning phase illustrated in Fig. 1 *C* and *F*, we generated options for the pointing and the mental arithmetic tasks that matched the classical options.

Participants did not receive feedback in the decision phase. Participants were instructed that one of their chosen options would be randomly selected at the end of the experiment (with values decreased by a factor of 10) and played for real money. Participants also knew that the probabilities used to generate the real outcome would be based on their own ability in the learning sessions.

How Good Were People's Choices?

The perception–cognition gap implies that people will make much better choices when they are faced with low-level decisions compared with when faced with classical decisions (or decisions based on high-level information). To evaluate how good participants' choices were, we used an approach typical in low-level studies (7): we compared participants' actual choices to the choices of a hypothetical agent who always chooses the best option. The best option

‡Rigorously characterizing high- and low-level processes, systems, and tasks is difficult. In the limit, all high-level tasks incorporate some low-level components (and vice versa). We did not attempt a unique formal definition, but used tasks that are conventionally considered low or high level. Mental arithmetic is often assessed as part of IQ testing and thus considered a core cognitive ability. Pointing is a widely used perceptuo-motor task (13).

§A recording error resulted in the inclusion of warm-up trials for one participant in the arithmetic task. This inclusion resulted in a marginal underestimation of this participant's mental arithmetic ability (see Fig. S6). Excluding this participant from the analysis does not change the patterns observed across Figs. 3–5.

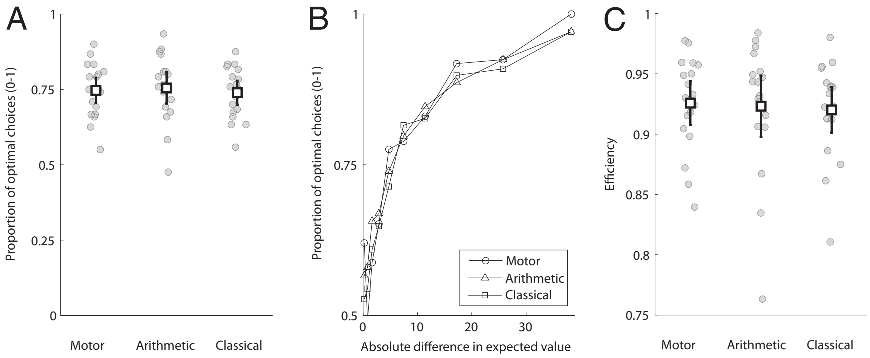

Fig. 3. Performance metrics as a function of task. (*A*) The proportion of choices maximizing expected value for each participant (gray discs) and the group average (black squares). Error bars are 95% CIs. (*B*) The proportion of choices maximizing expected value for each task, as a function of the absolute difference in expected value between pairs of choice options (discriminability), pooled across participants. (*C*) Efficiency for each participant (gray discs) and the group average (black squares). Error bars are 95% CIs. Gray discs representing individuals have been jittered laterally.

is the option that gives the highest reward in the long run (the option with the highest expected value).[¶]

Fig. 3*A* shows the proportion of optimal choices both for individual participants (gray discs) and the average proportion of optimal choices for each task (black squares). As can be seen, the average proportion of choices that maximized expected value was ~0.75, regardless of task. Note also that there were large individual differences.

An average optimal choice rate of 0.75 is moderately impressive. However, expected value maximization implies that the decision-maker is able to perfectly discriminate between choice options. Perfect discrimination is likely to be unattainable. Indeed, previous studies have found that choice consistency increases when choice options become more discriminable (23) and that expected value maximization better describes choice when options are easy to discriminate (24, 25).

As Fig. 3*B* shows, the proportion of optimal choices increased as the options became more discriminable (i.e., the absolute difference in expected value increased). Increasing differences between choice options do not only affect people's ability to discriminate, they are also related to the potential loss of making a suboptimal choice. Choices that are hard to discriminate are approximately equally valuable. Consequently, the cost of picking the suboptimal choice is small. Conversely, when choices are easy to discriminate, it is because one option is clearly more valuable than the other. Thus, for easily discriminated options, the cost of choosing the suboptimal option can be very high. If participants choose the wrong option mainly when options are hard to discriminate, Fig. 3*A* may give an overly pessimistic picture of participants' choice performance.

Fig. 3*C* shows our participants' expected earnings relative to the earnings of a hypothetical participant who always chooses optimally. Efficiency, or actual earnings divided by the earnings achieved by an optimal participant, is the standard performance metric in low-level decision studies (7). As can be seen, the average participant is expected to earn ~92% of the optimal earnings (with some expected to earn closer to 98%). Thus, whatever choice strategies our participants used, their strategies were nearly as efficient as the optimal one.[‖]

Fig. 3*C* also suggests that there were practically no differences in performance across the three tasks. For statistical comparisons when the null hypothesis is of interest, Bayesian tests are appropriate (27, 28). A comparison using JZSBayes factors (28) supports the hypotheses of equal efficiency levels across the three tasks [cognitive vs. pointing: Bayes factor = 5.59, $t(17) = -0.249$, $P = 0.81$; cognitive vs. classical: Bayes factor = 5.3, $t(17) = 0.335$, $P = 0.74$; pointing vs. classical: Bayes factor = 3.56, $t(17) = 0.979$, $P = 0.34$].[**] Thus, we found evidence against the perception–cognition gap. Performance was equally good across low- and high-level decisions.

Do People Deviate Systematically from Optimality?

Of course, the fact that decisions were highly efficient does not necessarily mean that they were free from bias, or, for that matter, that any biases were identical across tasks. Were there differences in choice strategies despite the comparable performance? Specifically, were low probabilities treated differently across the three tasks? Given the description–experience gap, we would expect our participants to act as if low probabilities are underweighted in the two tasks that rely on experience (pointing and mental arithmetic) but to act is if probabilities are overweighted in the classical task.[††]

Any decision model that allows for under- and overweighting of low probabilities could in principle be used to test for differences in how participants treat unlikely events. We fit cumulative prospect theory (5), a model commonly used to account for deviations from optimal decision-making. Two key aspects of the theory are its value and probability[‡‡] weighting functions. These functions allow values and probabilities to deviate systematically from objective ones. That is, they allow for the possibility that people assign too much, or too little, weight to values and probabilities. This flexibility in turn allows the model to account for many of the empirically observed deviations from optimal choice in classical studies. To make comparisons robust we tested a broad range of weighting functions used in the literature (*SI Text*, *Model Fitting*, Table S1, and Fig. S1).

Fig. 4 shows the best fit value and probability weighting functions for each task and participant. The panels show objective values (*A*–*C*) and probabilities (*D*–*F*) plotted against weighted values and probabilities. If participants made decisions using the objective values and probabilities, the best fitting function for each participant (Fig. 4, gray lines) would lie along the diagonal. Average value functions (Fig. 4, dashed lines, top row) below the identity line, as in *A*–*C*, are well-established phenomena likely to reflect the fact that the perceived value of money is such that the utility of £2000 is less than twice as much as £1000 (diminishing marginal utility).

For the best fit probability weighting functions (Fig. 4, bottom row), two trends are noteworthy. First, the average functions (dashed

[¶]Expected utility theory is a more liberal, and in the cognitive literature more commonly applied, standard. However, as low-level studies tend to use the more conservative standard of expected value maximization and find good agreement between it and participants' behavior, we used it here also. Note that expected value maximization can be argued to be the proper normative standard in typical behavioral experiments (21, 22).

[‖]Although absolute efficiencies across studies should be compared with care, the efficiencies obtained are comparable with some previous studies showing near-optimal low-level performance (e.g., mean efficiency: ref. 26 = 0.88, ref. 7 = 1.01).

[**]An anonymous reviewer suggested that the highly similar performance might hold only for easily discriminable choice options. Repeating the analysis with option pairs of expected value ratios <2 (as in ref. 25) produces near identical results to the ones reported above.

[††]The precise mechanisms for the observed differences in choice patterns across the two paradigms appear unclear. Compare, for example, refs. 16 and 29 and see refs. 14 and 15 for lists of possible mechanisms. We are primarily interested in detecting potential differences in the way people appear to treat low probabilities and less in determining the precise mechanisms behind such differences.

[‡‡]Probability weights are often termed "decision weights" to emphasize that probabilities are weighted in the decision process and that the weighting function does not represent a bias in probability judgment (3).

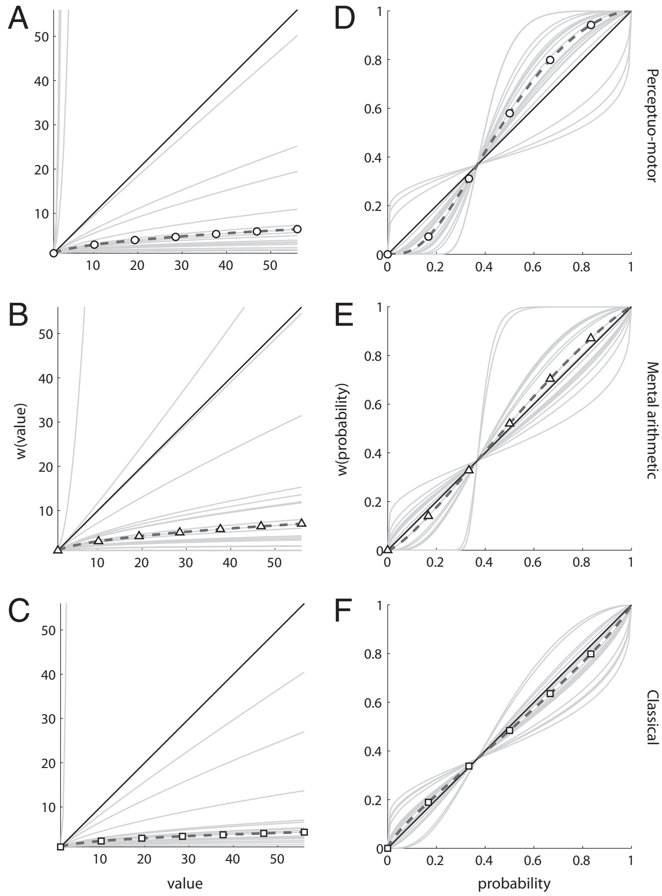

Fig. 4. Best fit value and probability weighting functions. (*A*–*C*) Participants' best fit value functions (gray lines) and group averages (black dashed lines). (*D*–*F*) Participants' best fit probability functions (gray lines) and group averages (black dashed lines). Functions (gray lines) below the diagonal represent underweighting of objective values and probabilities. Functions above the diagonal represent overweighting of objective values and probabilities. Functions on the diagonal represent objective values and probabilities. Group averages are exponents of the 50% trimmed mean on the logarithm of individual weights. This method takes into account the fact that weights <1 and weights >1 have different ranges (0–1 and 1–∞). Circles, perceptuo-motor; triangles, mental arithmetic; squares, classical.

lines) suggest underweighting of low probabilities for the pointing and arithmetic task but suggest overweighting for the classical task. The results thus appear to replicate the description–experience gap (14, 15)[§§] and to replicate previous results comparing low-level decisions to classical ones (20, 34, but see ref. 35). All three of the average probability weighting parameters can be shown to deviate from linear weighting (*SI Text*, *Statistical Tests on the Probability Weighting Parameters*, and Fig. S2). However, statistical significance does not also imply importance, an issue we turn to below.

The second noteworthy trend is that, consistent with studies that look at individual differences (20, 34, 36), parameter estimates were highly variable. Some participants appear near optimal with near-linear probability weights, whereas others show severe under- or overweighting. Specifically, for neither decision from experience nor for decisions from description is there a consistent pattern of only under- or overweighting of low probabilities.

To explore the relative importance of the average between-task trends and the large within-task variability (Fig. 4), we contrasted two ways of predicting people's choices. It is possible to predict a given participant's choices in a given task on the basis of (*i*) their choices in other tasks and (*ii*) other participants' choices in the same task. If the variability is caused by stable individual preferences and less by the type of task one is engaged in, one might expect between-task predictions (*i*) to be at least as good as within-task predictions (*ii*). If, on the other hand, different tasks give rise to wildly different choice behaviors, within-task predictions (*ii*) should be superior.

To test the usefulness of these two methods for predicting participants' choices, we used a permutation test (*SI Text*, *Within Tasks vs. Between Tasks*, and Fig. S3). Briefly, we first computed how well one can predict participants' choices with both methods. Each of the two resulting fit metrics was compared with its own null distribution. Distributions were generated by permuting the mapping between subjects and tasks. The two distributions describe, respectively, (*i*) the hypothesis that there is no mapping between individual participants' choices in one task and their choices in the other tasks and (*ii*) the hypothesis that there is no mapping between tasks and group choices.

The predictability of the between-task within-subject method was significantly better than chance ($P < 0.05$, two-tailed). For the within-task between-subject method, we failed to reject the null hypothesis ($P > 0.05$, two-tailed). In other words, knowing participants' choices in some tasks is useful for predicting participants' choices in other tasks. However, knowing the group's choices in one task seem no more useful than knowing the group's choices in unrelated tasks. This result implies that the variability across tasks was in large part due to stable individual differences and that the apparent group differences across tasks shown in Fig. 4 have limited predictive power.

Do People Deviate from Optimality when Biases Are Taken into Account?

The description–experience gap is typically evaluated under the assumption that objective probabilities correspond to subjective probabilities (16, 29, 30),[¶¶] and perceptuo-motor studies have typically assumed the same (20, 34, 37). Sometimes this assumption is tested (29), and other times it is not (30). If subjective probabilities are not calibrated, this and not underlying changes in preferences may underlie the gaps. We show in *SI Text Calibration of Subjective Probabilities*, and Figs. S4–S6 that participants' probability estimates were indeed not perfectly calibrated. It seems sensible then to check whether the moderate differences we observed for the average probability weighting functions in Fig. 4 can be accounted for by these discrepancies between subjective and objective probabilities. To this end, we repeated the model fitting process, replacing the objective probabilities for each choice option with probabilities estimated from ratings participants provided in the learning sessions (*SI Text*, *Calibration of Subjective Probabilities*).

When subjective rather than objective probabilities are fit, the results change dramatically (Fig. 5). Now, most participants appear to overweight small probabilities, just as they do when probabilities are given explicitly (classical task). Note also that the value weights (Fig. 5, top row) show the same qualitative pattern as in Fig. 4 (the same diminishing marginal utility across the three tasks).

[§§]Note that, although people choosing as if they underweight rare events is considered a characteristic of decisions from experience (14, 15), studies that fit (cumulative) prospect theory to evaluate underweighting, instead of inferring underweighting directly from peoples' choice patterns (16), have yielded mixed findings. Although some studies find probability weighting parameters suggestive of underweighting (29, 30), others find weights suggesting near-linear weighting (31), and others find overweighting (32, 33). Without systematic empirical investigation, it is difficult to trace the origin of these differences. The studies often differ substantially both in design and analysis. However, here we use a large number of different choice options, fit individual choice data in a setting where expected value learning strategies have been ruled out (*Discussion*), and use a maximum-likelihood approach that essentially avoids the problem of flat maxima (29).

[¶¶]Subjective probabilities are participants' beliefs about objective probabilities. Objective probabilities are the probabilities participants have beliefs about. Here, objective refers to participants' actual ability to hit targets, as measured by their task performance. Subjective refers to participants' beliefs about their ability to hit targets, as measured by participants' explicit ratings. This distinction is also sometimes made in the experience literature (29). The distinction is orthogonal to another distinction between experienced (the distribution participants actually experience) and objective probabilities (the distribution participants would experience if they saw infinitely many samples) in the same literature (30). Experienced probabilities using this distinction corresponds to objective probabilities as defined here.

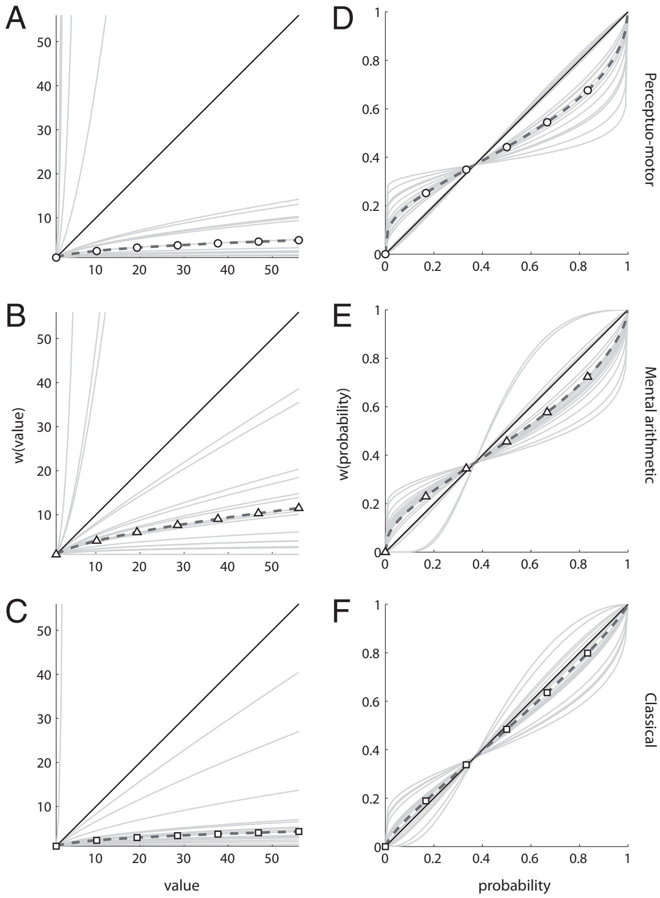

Fig. 5. Best fit value and probability weighting functions: subjective probabilities. (*A–C*) Participants' best fit value functions (gray lines) and group averages (black dashed lines). (*D–F*) Participants' best fit probability functions (gray lines) and group averages (black dashed lines). Functions (gray lines) below the diagonal represent underweighting of objective values and probabilities. Functions above the diagonal represent overweighting of objective values and probabilities. Functions on the diagonal represent objective values and probabilities. Circles, perceptuo-motor; triangles, mental arithmetic; squares, classical.

The above fits, and the differences in fits across subjective and objective probabilities generalize beyond the specific parameterization of cumulative prospect theory reported here (Fig. S1). Furthermore, as shown in Fig. S7, a simple model, assuming expected value maximization and less-than-perfect discrimination, generally accounts better for the data when subjective probabilities are fit compared with when objective probabilities are fit. This result suggests that the reversals in probability weighting functions are not likely due to participants simply being bad at reporting their beliefs.

Discussion

Once decision-making tasks are equated, numerical (classical), high-level (mental arithmetic), and low-level (perceptuo-motor) choices seem equally good, challenging the idea of a perception–cognition gap (see ref. 38 for converging evidence from the time domain). Moreover, although a model allowing probabilities to be weighted differently across tasks replicates typically found underweighting for the experience tasks and overweighting for the classical task when objective probabilities were fit, this contrast disappeared when subjective probabilities were fit, thus challenging the description–experience gap. What factors might explain the discrepancy between our results and previous results in support of decision-making gaps?

Arguably, the most parsimonious explanation for the perception–cognition gap lies in the very different methods by which performance is assessed. Indeed, when we used the standard low-level methods, that of evaluating actual choice performance, we found that, although participants were not precisely optimal, they were nevertheless highly efficient, regardless of task. In contrast, when we fit cumulative prospect theory, many participants had probability weights that were substantially different from 1, suggesting that they weight probabilities suboptimally, regardless of task. Thus, we were able to show both good performance (using standard low-level metrics) and bad performance (using standard high-level metrics) for both low- and high-level decisions.

The explanation for the absence of a description–experience gap is perhaps less clear. We replicated a key qualitative difference when fitting objective probabilities but not when fitting subjective probabilities. This result implies that the gap has arisen because it is erroneously assumed that subjective probabilities match objective ones. However, there are several caveats to this conclusion. First, subjective probabilities can be elicited in many ways. The method used here has some advantages, but is not perfect (*SI Text, Calibration of Subjective Probabilities*). Future studies might explore different elicitation methods or perhaps entirely avoid explicit estimates by relying on pairwise more-likely/less-likely judgments across domains. Regardless of the merits of different elicitation methods, the finding of equal efficiency across description and experience implies an upper bound on the description-experience gap. That is, even if there is something fundamentally different about our decisions when we rely on experience, that something does not appear to influence the efficiency with which decisions are made.

A second caveat is that our tasks are different from standard experience tasks. Our participants did not, for example, sample from binomial distributions. Instead, they learned about their own skills and used this knowledge to make decisions. Therefore, explanations for the description–experience gap, based on, for example, sampling error or recency effects (14, 15), are not directly applicable. Because of such differences, it is presently unclear whether our results generalize to more commonly used experience-based tasks. That said, there is one crucial point to be made in this regard.

The present tasks are different from standard experience tasks, partly because many of the existing confounds between description and experience tasks were eliminated. For example, in the classical paradigm and in our tasks, one has to somehow trade off probabilities with values to reach a decision. It is possible that people also do this in experience-based tasks. However, in these tasks, a much simpler strategy is available. Participants could learn how much money each button returns on average (29, 31), much like rats can learn which of two options to prefer (39). Such a learning strategy would mean that a direct tradeoff between probabilities and values is not needed and that the task reduces to a choice between two expected values. Because we, to a large extent, eliminated such confounds, our study arguably represents a stronger test of the hypothesis that there are fundamental and theoretically important differences in how people make decisions when relying on their own experiences vs. when relying on descriptions.

Despite equally good performance across tasks, the same apparent suboptimal weighting of probabilities across tasks, and the relative stability of preferences across time (*SI Text, Are Preferences Stable Across Time?*), our results do not imply that decisions made using different modalities, or on the basis of different kinds of information, are identical. On average, 22% of our participants' choices were different across task pairs. That is, our participants unsurprisingly did not make identical choices across the three tasks. Participants also made perceptuo-motor decisions faster than they made either mental arithmetic or classical decisions (Fig. S8), perhaps because it is easier to discriminate probability information in the guise of perceptuo-motor targets. Whatever the cause, the fact that perceptuo-motor decisions were faster suggests that the processes underlying the different decisions are not identical. It is thus of fundamental importance that the perceptuo-motor literature and the experience literature have hugely broadened the way we study decision-making. Nevertheless, our results question whether the aforementioned differences merit the distinction gap.

In the most important respect, that of potential earnings, there was no difference across the tasks, challenging the notion of gaps. Moreover, the small task-related differences we did observe were dominated by consistent individual differences. Importantly, because efficiencies were high, the ways in which people do deviate from optimal decision-making do not seem particularly costly. Even classical decisions were surprisingly good, a result that is all the more important because it was obtained across a large, random sample of choice options (40), and not, as has been typical in cognitive studies, with select choice options. Put another way, people may not be perfectly rational, but their irrationalities do not seem to lead them far astray from optimality, not just with low-level decisions but even in classical description-based tasks.

Methods

Eighteen members of the local participant panel were paid £6/h (plus a possible bonus of £0–£6) to participate (total participation time, ~2.5 h). Informed consent was obtained, and the study was approved by the School's Ethics Committee. A tablet (Wacom DTZ-2100) was used for the pointing task. Experiments were written in Matlab using Psychtoolbox (41, 42).

The two learning sessions were counterbalanced across participants. There were 120 option pairs for each format in the decision-making stage. The options were presented in a randomized order with no imposed time limit. Unlike most previous studies (3, 20, 31, but see ref. 43), we used randomly selected choice options to achieve a wide range of differences in expected values. Probabilities were drawn from the range 0.05–0.95 (in steps of 0.05). Values were drawn from £1 and £3 to £54 (in steps of £3). Five of 120 pairs were chosen so that one option dominated the other (i.e., both probability and value was higher for one option). Most participants chose the dominating option most of the time (mean number out of 5 = 4.85, minimum = 4, maximum = 5), indicating that participants paid attention and understood the decision task.

In fitting cumulative prospect theory, a standard probability weighting function (44), value function (5), and a logistic choice function (45) were used. The latter captures the fact that people are not perfectly sensitive to differences between options (Fig. 3*B*). Model parameters were determined using maximum likelihood estimation, fitting separately for each participant and task. Unlike most studies, we tested several model parameterizations to ensure fits were not due to the selection of a particular parameterization. For details on the model fitting and this model exploration, see *SI Text, Model Fitting*.

ACKNOWLEDGMENTS. A.J. was supported by the Economic and Social Research Council UK.

1. Allais M (1952/1979) The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American school. *Expected Utility and the Allais Paradox*, eds Allais M, Hagen O (Springer, New York), pp 227–145.
2. Ellsberg D (1961) Risk, ambiguity and the savage axioms. *Q J Econ* 75(4):643–669.
3. Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–292.
4. Tversky A, Kahneman D (1981) The framing of decisions and the psychology of choice. *Science* 211(4481):453–458.
5. Tversky A, Kahneman D (1992) Advances in prospect theory: Cumulative representation of uncertainty. *J Risk Uncertain* 5(4):297–323.
6. Trommershäuser J, Maloney LT, Landy MS (2003a) Statistical decision theory and the selection of rapid, goal-directed movements. *J Opt Soc Am A Opt Image Sci Vis* 20(7):1419–1433.
7. Trommershäuser J, Maloney LT, Landy MS (2003b) Statistical decision theory and trade-offs in the control of motor response. *Spat Vis* 16(3–4):255–275.
8. Whiteley L, Sahani M (2008) Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. *J Vis* 8(3):1–15.
9. Navalpakkam V, Koch C, Perona P (2009) Homo economicus in visual search. *J Vis* 9(1):1–16.
10. Navalpakkam V, Koch C, Rangel A, Perona P (2010) Optimal reward harvesting in complex perceptual environments. *Proc Natl Acad Sci USA* 107(11):5232–5237.
11. Warren PA, Graf EW, Champion RA, Maloney LT (2012) Visual extrapolation under risk: Human observers estimate and compensate for exogenous uncertainty. *Proc Roy Soc B Biol Sci* 279(1736):2171–2179.
12. Trommershäuser J, Landy MS, Maloney LT (2006) Humans rapidly estimate expected gain in movement planning. *Psychol Sci* 17(11):981–988.
13. Trommershäuser J, Maloney LT, Landy MS (2008) Decision making, movement planning and statistical decision theory. *Trends Cogn Sci* 12(8):291–297.
14. Hertwig R, Erev I (2009) The description-experience gap in risky choice. *Trends Cogn Sci* 13(12):517–523.
15. Rakow T, Newell B (2010) Degrees of uncertainty: An overview and framework for future research on experience-based choice. *J Behav Decis Making* 23(3):1–14.
16. Hertwig R, Barron G, Weber EU, Erev I (2004) Decisions from experience and the effect of rare events in risky choice. *Psychol Sci* 15(8):534–539.
17. Hertwig R, Pleskac TJ (2010) Decisions from experience: Why small samples? *Cognition* 115(2):225–237.
18. Maloney LT, Trommershäuser J, Landy MS (2007) Questions without words: A comparison between decision making under risk and movement planning under risk. *Integrated Models of Cognitive Systems,* ed Gray W (Oxford Univ Press, New York), pp 297–313.
19. Fox CR, Tversky A (1998) A belief-based account of decision under uncertainty. *Manage Sci* 44(7):879–895.
20. Wu S-W, Delgado MR, Maloney LT (2009) Economic decision-making compared with an equivalent motor task. *Proc Natl Acad Sci USA* 106(15):6088–6093.
21. Rabin M (2000) Risk aversion and expected-utility theory: A calibration theorem. *Econometrica* 68(5):1281–1292.
22. Rabin M, Thaler RH (2001) Anomalies: Risk aversion. *J Econ Perspect* 15(1):219–232.
23. Mosteller F, Nogee F (1951) An experimental measurement of utility. *J Polit Econ* 59(5):371–404.
24. Brandstätter E, Gigerenzer G, Hertwig R (2006) The priority heuristic: Making choices without trade-offs. *Psychol Rev* 113(2):409–432.
25. Brandstätter E, Gigerenzer G, Hertwig R (2008) Risky choice with heuristics: Reply to Birnbaum (2008), Johnson, Schulte-Mecklenbeck, and Willemsen (2008), and Rieger and Wang (2008). *Psychol Rev* 115(1):281–290.
26. Gepshtein S, Seydell A, Trommershäuser J (2007) Optimality of human movement under natural variations of visual-motor uncertainty. *J Vis* 7(5):1–18.
27. Gallistel CR (2009) The importance of proving the null. *Psychol Rev* 116(2):439–453.
28. Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G (2009) Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev* 16(2):225–237.
29. Ungemach C, Chater N, Stewart N (2009) Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychol Sci* 20(4):473–479.
30. Camilleri AR, Newell BR (2011) When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychon Bull Rev* 18(2):377–384.
31. Hau R, Pleskac TJ, Kiefer J, Hertwig R (2008) The description-experience gap in risky choice: The role of sample size and experienced probabilities. *J Behav Decis Making* 21(5):493–518.
32. Fox CR, Hadar L (2006) Decisions from experience = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004). *Judgm Decis Mak* 1(2):159–161.
33. Abdellaoui M, L'Haridon O, Paraschiv C (2011) Experienced vs. described uncertainty: Do we need two prospect theory specifications. *Manage Sci* 57(10):1879–1895.
34. Wu S-W, Delgado MR, Maloney LT (2011) The neural correlates of subjective utility of monetary outcome and probability weight in economic and in motor decision under risk. *J Neurosci* 31(24):8822–8831.
35. Glaser C, Trommershäuser J, Mamassian P, Maloney LT (2012) Comparison of the distortion of probability information in decision under risk and an equivalent visual task. *Psychol Sci* 23(4):419–426.
36. Gonzalez R, Wu G (1999) On the shape of the probability weighting function. *Cognit Psychol* 38(1):129–166.
37. Nagengast AJ, Braun DA, Wolpert DM (2011) Risk-sensitivity and the mean-variance tradeoff decision-making in sensorimotor control. *Proc Roy Soc B Biol* 278(1716):2325–2332.
38. Jarvstad A, Rushton SK, Warren PA, Hahn U (2012) Knowing when to move on: Cognitive and perceptual decisions in time. *Psychol Sci* 23(6):589–597.
39. Thorndike EL (1911) *Animal Intelligence* (Macmillan, New York).
40. Dhami MK, Hertwig R, Hoffrage U (2004) The role of representative design in an ecological approach to cognition. *Psychol Bull* 130(6):959–988.
41. Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10(4):433–436.
42. Pelli DG (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat Vis* 10(4):437–442.
43. Erev I, Roth AE, Slonim R, Barron G (2002) Combining a theoretical prediction with experimental evidence. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1111712. Accessed August 5, 2010.
44. Prelec D (1998) The probability weighting function. *Econometrica* 66(3):497–527.
45. Stott HP (2006) Cumulative prospect theory's functional menagerie. *J Risk Uncertain* 32(2):101–130.