Research Article

Title: A Tale of Two Environments: Divisive Normalization and the (In)Flexibility of Choice

Short title: Divisive Normalization and Flexibility of Choice

Vered Kurtz-David[a]*, Shreya Sinha[a], Vinayak Alladi[b,c], Stefan Bucher[d], Adam Brandenburger[e], Kenway Louie[a,f], Paul Glimcher[a,e] # and Agnieszka Tymula[b] #.

[a] Neuroscience Institute, Grossman School of Medicine, NYU Langone, 435 E 30th St., New York, NY 10016 USA.
[b] School of Economics, University of Sydney, A02 – Social Sciences Building Sydney, NSW 2006 Australia.
[c] The World Bank Group, 1818 H St NW, Washington, DC 20433 USA.
[d] Sloan School of Management, MIT, Cambridge, MA 02142 USA and Max Planck Institute for Biological Cybernetics, Maria-von-Linden-Straße 6, Tübingen 72076 Germany.
[e] NYU Stern School of Business, NYU Tandon School of Engineering, and NYU Shanghai, New York University, 44 West 4th Street, New York, NY 10012, USA.
[f] Center for Neural Science, NYU, 4 Washington Place, New York, NY 10003 USA.

*Corresponding author
Email: kurtzv02@nyu.edu

# equally contributing senior authors.

Keywords: decision-making, efficient coding, divisive normalization, risky choice

This PDF file includes:

Main Text
Figures 1 to 5
Tables 1 to 2

Supporting Materials
Figures S1 to S3
Tables S1 to S5

Abstract

The Divisive Normalization (DN) function has been described as a “canonical neural computation” in the brain that achieves efficient representations of sensory and choice stimuli. Recent work indicates that it efficiently encodes a specific class of Pareto-distributed stimuli. Does the brain shift to different encoding functions or is there evidence for DN encoding in other types of environments? In this paper, using a within-subject choice experiment, we show evidence of the latter. Subjects made decisions in two distinct choice environments with choice sets drawn from either a Pareto distribution or a uniform distribution. Our results indicate that subjects' choices are better described by a divisive coding strategy in both environments. Moreover, subjects appeared to calibrate a DN function to match, as closely as possible, the actual statistical properties of each environment. These results suggest that divisive representations of encoded stimuli may be inherent to the nervous system.

Teaser

Humans use the same value encoding function across statistical environments, suggesting a fixed feature in human cognition.

Introduction

We make some decisions more often than others – in dozens of instances during our life we choose between having a pizza or a burger for dinner, but rarely have to indicate which of the two-starred Michelin restaurants we prefer. An often overlooked fact is that these encounter frequencies play a critical role in defining efficient encoding strategies – given constraints on neural coding, more accurate encoding must generally be allocated to more frequently encountered stimuli (*1*, *2*). Indeed, experimental studies confirm this theoretical insight, showing a dependency of preference orderings, choice patterns (*3*–*7*), and choice efficiency (*3*, *8*) on the frequency with which subjects encounter different rewards.

This has led to the conclusion that during the decision process, the brain adheres to principles of *efficient coding*, economizing the allocation of resources to optimize decision outcomes (*3*, *8*–*13*). A canonical example of a well-studied efficient code (*13*, *14*) is Divisive Normalization (henceforth DN) (*15*), which has been related to neuronal firing rates across all sensory modalities (*16*–*19*) and across various cognitive domains as well (*20*). The DN function enables a system with limited information capacity to employ a flexible and scale-invariant encoding of naturally occurring stimuli that is sensitive to encounter frequency (*17*, *21*, *22*). Ample evidence has supported the notion that DN is also highly predictive of reward value encoding in the human and animal choice mechanism (*6*, *7*, *23*–*25*).

At least one form of DN has been analytically shown to be an efficient code for stimuli with a probability of occurrence that is described by the asymmetrical long-tailed Pareto Type III distribution (see eq. (iv-vi) in Materials and Methods) (*26*). This prompts the empirical question of whether the brain employs non-DN encoding functions when the statistical properties of the input stimuli (in our case, *choice environment*s) are not Pareto-distributed. Would we expect to find evidence of divisive encoding mechanisms (*13*) – like cross-normalized DN (*26*) – only in Pareto-distributed environments? If so, this might imply that previous documentation of DN encoding mechanisms may say more about the antecedent stimulus distributions used in experiments than about constraints on encoding. An alternative hypothesis, however, is that our brains are constrained to employ DN-like encoding mechanisms (*13*). Such a constraint might reflect an adaptation of the nervous system to Pareto-distributed real-world natural stimuli, such as the sensory (*14*, *18*, *27*) and ecological (*28*, *29*) environments we typically encounter.

In this study, we present our subjects with a binary-choice task in two environments characterized by different reward distributions. In one environment, valuations are Pareto Type III distributed and hence at least some DN functions are efficient (*13*, *26*). In the other environment, valuations are uniformly distributed. In such an environment, a DN encoder would not represent the environment most efficiently (*26*) (Fig. 1A). We test hypotheses about our subjects' value encoding functions by fitting the patterns of errors in their choices with two random utility models (henceforth, RUM) (*30*, *31*). The first function captures the key features of the family of DN models (*32*, *33*). The second function is a RUM with a power utility function that nests within its parametric family a linear, a concave, and convex encoder (henceforth, power utility; Fig. 1B). Power utility is the standard model in economic research and lies outside the DN family. We test whether subjects are better described as obligate-DN choosers who use DN in both environments or whether subjects' choices are better described with a DN function in one environment and with a power utility function in the other (Fig. 1C).

In line with theory (*13*, *26*), we find that in a Pareto-distributed environment, subjects employ an encoding of values that is well modeled by DN. However, we also find that the DN model better captures subjects' choices in the uniformly-distributed environment. This suggests that subjects' choices are more accurately described by divisive encoders, such as those found in DN models, than by standard power utility functions. We also find evidence for context dependency in subjects' choices in the sense that, within the constraints of DN encoding, subjects partially adapt to changes in the specific statistical properties of the choice environment.

Taken together, our results suggest that divisive normalization may be an inherent component of the encoding mechanism used during the choice process. Future work could generalize these findings to other types of statistical environments. Finally, the current study focuses on decision making processes, but, given the dominance of DN representations across cortical systems, our findings may be of general interest to the study of encoding mechanisms in sensory and other cognitive domains.

[Insert Fig. 1 here]

Results

Two-Stage Task Design

Seventy-six subjects completed a two-stage choice task. In STAGE I (Fig. 2A left panel), subjects reported their valuations (willingness to pay) for 33 50-50 lotteries that pay either y_1 or y_2 dollars (see Table S1 for a complete lottery list). These valuations were used to estimate, for every subject i, their STAGE I subjective value function, $u_i = 0.5y_1^{\rho_i} + 0.5y_2^{\rho_i}$, using a standard non-linear least squares (NLS) estimation. The subjective value function curvature (ρ_i) varied substantially from subject to subject (Fig. 2E). Using individual ρ_i estimates, we generated subject-specific distributions of rewards in terms of their subjective – rather than dollar – values for the STAGE II task (Fig. 2B). This first step was critical. It allowed us to perform all our analyses in the domain of subjective value, removing simple utility curvature from our primary analyses and allowing us to create individualized choice sets with specific distributional properties that were essential for our design. Without this transformation, small subject-specific differences in utility curvature (risk attitudes) would have made the construction of probative choice sets required for the experiment impossible.

In STAGE II, on two separate days, subjects made binary choices between 50-50 lotteries (Fig. 2A, right panel), with 320 decisions on each day. We created two choice environments: on one day, subjects were choosing between lotteries with subjective values drawn from a Pareto Type III distribution, for which DN has been proven to be an efficient encoder (henceforth, Pareto), and on the other day between lotteries with subjective values drawn from a uniform distribution for which DN has been shown to not be an efficient encoder (*26*). Subjects encountered each distributional environment on a different day (counter-balanced across subjects) to avoid contextual spillovers.

Using these risky-choice lotteries, rather than choices over consumer goods, enabled us to generate *continuous* distributions of valuations for STAGE II and to fully control their *distributional shape*. As noted above, our decision to generate the distributions of STAGE II lotteries in subjective value space, rather than in dollar space, aimed to control for the heterogeneity in subjects' subjective valuations of lotteries (risk attitudes, Fig. 2E and Table S2). If, instead, we had created these distributions based on the dollar amounts, then for most subjects (all those with a non-linear subjective value function, i.e. $\rho_i \neq 1$) STAGE II choice sets would have not corresponded to uniform or Pareto type III distributions due to the heterogeneity in their subjective valuation of monetary lotteries. Our two-stage procedure thus ensured that for all subjects the environments had the same distributional shape – enabling us to assess the effect of the distributional properties of the two choice environments, while fully controlling for individual differences in preferences.

Across subjects, we fixed the first moment (mean) of valuations and the range of monetary payoffs in both environments. Naturally, the second moment (standard deviation) of the uniform distribution was also fixed across subjects. The second moment of the Pareto distribution (as measured in dollars) varied by subjects' subjective valuations of money as assessed in STAGE I (risk attitudes) (Fig. 2C, Fig. S1).

Accordingly, this heterogeneity also varied the distributions of the high and low monetary payoffs in each lottery (Fig. S1). To ensure that we fully captured each distributional environment, we matched the mean and standard deviation of the choice sets with those of larger sets of 100k draws (Fig. 2D). See Materials and Methods for further details on our sampling design.

Overall, subjects appeared to pay careful attention during the study – only six subjects in the uniform environment, and nineteen subjects in the Pareto environment failed to choose the higher subjective value lottery in more than 20% of trials (Fig. S2A). Respectively, on average, subjects violated first-order stochastic dominance in 0.97% of trials in the uniform treatment and in 1.08% of trials in the Pareto treatment (Fig. S2B). Note that a higher incidence of mistakes in the Pareto environment is expected – as due to the correlational structure across lotteries, the value difference between lotteries was (on average) smaller and thus choices were harder in the Pareto environment(*34*). Finally, even though the experiment was quite demanding (320 trials in each of the two sessions), subjects' performance was not affected by fatigue – the propensity to choose the lottery with the higher subjective value did not vary between the first and second halves of each experimental session (Pareto sessions: p=0.2791, uniform sessions: p=0.5109, paired t-test (df=75), Fig. S2C).

[Insert Figure 2 here]

Distributional Properties of the Choice Environments Influence Subjects' Choice Behavior

Our overarching goal was to study how the distributional properties of the choice environment influenced the encoding of value, and whether subjects could flexibly switch between different types of encoding mechanisms, as evidenced by errors in their choice patterns, in different environments. We created the experimental choice environments with Pareto Type III and uniform distributions of valuations. In this section, we tackle the first part of our research question in a model-free manner, determining whether the distributional structure influenced the errors produced by our subjects in a meaningful manner.

It is useful to introduce our hypotheses using an illustration. In Fig. 3A-B, we indicate the probability of choosing lottery 1 with valuation u_1, given the coupling of the (u_1, u_2) valuations in a choice set. Choices along the diagonal represent trials in which the two lotteries had the same or very similar valuations, whereas trials that are away from the diagonal correspond to choice sets in which the two lotteries' valuations were substantially different. A central feature of DN is the calibration of the function to the input stimuli. That is, resources are allocated to the range of stimuli most likely to be observed (*tuning*) (*15*, *25*). Thus, compared to non-divisive encoders, if DN governs the choice mechanism in a Pareto environment, we would expect that in this environment subjects would make more mistakes in choice sets with elements away from the high-density center of the main diagonal, because these choices are less frequent.

Conversely, we also expect that subjects in the Pareto environment would make fewer mistakes in choices whose valuations lie near the main diagonal because these near-equivalued choices occur more frequently. We find both patterns in our data.

To statistically test whether the frequency of mistakes increased faster as choice sets moved away from the main diagonal in the Pareto environment than in the uniform environment, we ran a probit regression with an indicator dependent variable equal to one for trials on which a subject selected the option with higher SV, and equal to zero otherwise. We controlled for the difference in difficulty across the trials by including the absolute value difference between the lottery valuations ($|u_1 - u_2|$) and for the general impact of the distribution by including a dummy for the Pareto distribution. The different rate of mistakes depending on the distance from the diagonal in each environment are captured by a significant coefficient on the interaction of Pareto dummy and ($|u_1 - u_2|$) (Column (1) in Table 1). As expected, we found that choice accuracies increased with an increase in the subjective value distance between the two options, and that moving from the uniform distribution to the Pareto distribution reduced accuracy (see also discussion in the previous section). Importantly, in line with our hypothesis, we found a negative and significant interaction term, indicating that relative to the uniform distribution, in the Pareto distribution, subjects were more likely to make errors once encountering choice sets further away from the diagonal, those sets that they experienced less often in the Pareto environment. We can thus conclude that encounter frequency as defined by the Pareto distributional structure did influence choice accuracy.

To examine whether subjects calibrated their encoding function to the most frequently presented choice sets, we tested if they made fewer mistakes around the high-density center of the main diagonal in the Pareto environment. We ran a complementary probit regression focusing on twenty-two valuation bins from the center of the distributions presented in Fig. 3A-B (out of an equally-spaced 40-bin space), which corresponded to lotteries with $9-42 payoffs (Column (2) in Table 1). The center (medians) of the distributions depended on subjects' subjective valuation of dollar amounts (ρ parameter). In the Pareto distribution, the smallest median was $11.45 and the highest was $33.58. Likewise, in the uniform distribution, the smallest median was $22.85 and the highest was $41.53. Thus, determining a range of $9-42 included the distributions' centers for all the subjects in our sample. See Table S3 for an alternative definition that included a smaller range.

In addition to the regressors used in the model above, we included a dummy variable that indicated whether a lottery was taken from around the diagonal of the valuation space, as well as its interaction with the Pareto distribution dummy. We defined lotteries as laying around the diagonal if the ratio between the two valuations was $0.9 < \frac{u_2}{u_1} < 1.1$.

Not surprisingly, choice accuracy was lower in choice sets around the diagonal, since these represented the most difficult choices in the experiment with the smallest SV difference. Crucially though, we found a positive interaction term between the diagonal and Pareto dummies, suggesting that relative to the uniform environment, in the Pareto environment, subjects had higher accuracy in those particularly difficult trials within the highly sampled region. Our results remained robust for other definitions of *the center of the distribution* and *around the diagonal* (Table S3).

Together, these results suggest that in the Pareto environment subjects adjusted their value encoding to increase choice accuracy rates at the center of the joint distribution, at the expense of the decreased choice accuracy at the margins. This is evidence for a divisive form of value encoding, where choice discriminability is the highest near the mode of the distribution (see Fig. 1B).

[Insert Table 1here]

Evidence for DN-like Value Encoding Across Choice Environments

The findings in the previous section provided initial evidence that subjects adapted to the distribution of valuations and that subjects used divisive encoding in the Pareto environment. Our next goal was to evaluate whether subjects used the same or different encoding mechanisms in each of the two environments. To answer this question, we tested which model, a generalized form of divisive normalization (DN) or power utility, better captures subjects' choices. We picked this DN model because it is regarded as a canonical encoding mechanism in the brain(*15*–*17*, *19*), including in the choice domain (*7*, *24*, *25*, *33*). Crucially, the DN model has been considered an efficient encoder (*1*, *9*, *13*, *35*), as at least one variant of the DN model has been proven to efficiently encode Pareto distributed environments(*26*). We thus expected some form of DN encoding in this environment. In the DN model, subject i's STAGE II subjective value function of a lottery $k \in \{1,2\}$ with payoffs $x_{1,k}$ or $x_{2,k}$ is given by:

$$\text{(i)} \quad S_{i,k} = 0.5\frac{(u_i(x_{1,k}))^{\alpha_i}}{(u_i(x_{1,k}))^{\alpha_i}+M_i^{\alpha_i}} + 0.5\frac{(u_i(x_{2,k}))^{\alpha_i}}{(u_i(x_{2,k}))^{\alpha_i}+M_i^{\alpha_i}} + \varepsilon_i$$

where α_i is the function's curvature, $u_i(\cdot)$ is subject i's STAGE I valuation (i.e., $u_i(x_{1,k}) = x_{1,k}^{\rho_i}$), M is the reward expectation, and ε_i is an additive decision noise drawn in each trial from a zero-mean normal distribution, such that $\varepsilon_i \sim N(0, \theta_{DN})$.

The second model we examined was the commonly used power utility model (*36*):

$$\text{(ii)} \quad R_{i,k} = 0.5\big(u_i(x_{1,k})\big)^{r_i} + 0.5\big(u_i(x_{2,k})\big)^{r_i} + \eta_i$$

The model has one free parameter (r), which captures the function's curvature. When $r = 1$, the function is linear. Similarly to our DN model, here, too, we included an additive decision noise $\eta_i \sim N(0, \theta_p)$.

For every subject, we estimated both models using maximum likelihood estimation (see Materials and Methods). The subject-specific recovered parameters are reported in Table S4, and the sample medians are in Table 2. To determine, at the population level, which model better captured subjects' choice patterns in each environment, we compared each subject's Bayesian Information Criterion (BIC) scores across the two models in each environment. Options in the uniform environment had on average higher value difference, thus responses in this environment were more accurate (Fig. S2) and less noisy (Fig. S3). Therefore, we only compare BIC scores of the two models within the same environment, and do not compare the models across the two environments.

In line with our hypothesis, we found that in the Pareto environment, subjects' BIC scores were on average significantly lower, indicating a better model fit, for the DN model than for the power utility model (Fig. 3C, one-sided Wilcoxon sign-rank test, Z=4.4603, p<0.0001). This was true for 48 subjects (out of 76). In the uniform environment, we expected that a model that does not belong to the family of DN models would better fit the data. Instead, we found that again the BIC scores were on average significantly lower for the DN model (Fig. 3D, one-sided Wilcoxon sign-rank test, Z=2.9692, p=0.0015) and this held for 42 (out of 76) subjects. Moreover, for only four subjects the curvature parameter in the power utility model was estimated as linear or as almost linear ($r = 1 \pm 0.05$). The mean and median r estimates were 0.608 and 0.366, respectively. Importantly, the asymmetrical distributions of the differences in BIC scores (see insets in Fig. 3C-D) indicate that while for most subjects both models do (almost) equally well, there is a group of subjects for whom the DN model predicts their choices much better (ΔBIC>20 for 29 subjects in Pareto and 18 subjects in uniform).

To further compare the two models, we estimated them on the sample level. Table S5 presents the recovered pooled estimates from this analysis. Note that this analysis could only be done in dollar-space to allow comparability of lotteries across subjects, and to recover meaningful estimates of the M parameter in the DN model. Here, too, we find that the DN model captured subjects' choices better, evident by the lower BIC scores when aggregating choices from both treatments (leftmost column), as well as within each environment (second and third columns). These results should be interpreted cautiously since the reward distributions were not fully controlled in the dollar space (Fig. S1).

Another way to examine the effect of the distributional environment on subjects' value-encoding – and validate our task design – is to examine the relationship between the estimates of subjects' subjective valuations of lotteries in STAGE I (ρ_i , see Fig. 2E) and STAGE II α parameter in DN, and r parameter in power utility. Due to the nature of our design, a hyperbolic relationship (i.e. $y = \frac{1}{x}$) would suggest that STAGE I curvature was undone in STAGE II in both models. In the power utility model, this would also imply linear encoding of monetary payoffs (because $(x^{\rho})^{\frac{1}{\rho}} = x$). In DN, it would mean that all curvature in STAGE II is associated with the DN encoding. Fig. 3E-G plots STAGE II parameters against STAGE I ρ.

In both Pareto and uniform environments, we saw a much closer hyperbolic relationship between ρ and the DN α parameter (Fig. 3 E and G) than between ρ and power utility r parameter (Fig. 3 F and H). We compared the root of mean squared errors (RMSE) between the hyperbolic function and the parameters in both models and confirmed that across the two environments, the α parameter of the DN model was more likely to maintain this hyperbolic relationship (α: Pareto: RMSE=0.6491, uniform: RMSE=0.6585; r: Pareto:

RMSE=0.8109, uniform: RMSE=0.8588). This suggests that the curvature of the STAGE II subjective value functions can be attributed to DN value coding.

Taken together, all these results strengthen the notion that subjects used DN encoding of value in both environments.

[Insert Table 2 here]

[Insert Fig. 3 here]

Context-Dependency: Adaptation of the Encoding Function to the Choice Environment

Our next aim was to examine whether subjects adapted their encoding according to the properties of the different environments. Another key difference between the uniform and Pareto environments was that, for all subjects in our sample, the medians of the subjective valuations in the uniform environments were higher than in Pareto (sample medians: 18.721 vs. 14.723 util units, respectively, Δ= 3.998, one-sided Wilcoxon sign-rank test between subject-specific medians, Z=7.572, p<0.0001). The reward expectation M in the DN model tracks the median of the reward distribution, and hence we hypothesized it would be higher in the uniform environment. Consistent with this hypothesis, the sample median of the recovered M parameters in the uniform environment was higher by 4.99 (in util units) than in the Pareto environment (Table 2, also corroborated by a one-sided Wilcoxon sign-rank test, Z=2.8907, p=0.0019). This difference between the recovered M parameters was very close to the actual difference between the distributions' medians, indicating that subjects – at least at the sample-level – quite precisely calibrated their encoding to the difference in reward expectation. On the subject level, we found that for 44 out of 76 subjects estimated M was higher in the uniform environment (Fig. 4A).

Our pooled estimation further supports this conclusion with $M(uniform)$=66.6531 and $M(Pareto)$=55.5609, second and third columns in Table S4, p<0.001. As a final robustness check, we estimated the DN model using the full dataset with the data from both environments, and included an additive dummy variable for the Pareto environment in the estimation of the M parameter ($M = constant + M_{Pareto} x Pareto$). The output of this model split M into a constant, corresponding to the estimate of M for the uniform environment, and M_{Pareto}, which captured the difference in M in the Pareto relative to the uniform environment. We found M_{Pareto} to be negative and significant (p<0.001), indicating M was lower in the Pareto environment.

In contrast to the M parameter, we had no prior hypotheses regarding the model's curvature parameter α. Nevertheless, comparing subject-specific estimates, we found that on average, the α parameter was higher by 0.1593 in the Pareto environment (one-sided Wilcoxon sign-rank test, Z=1.9987, p=0.0228, Fig. 4B). This result may indicate that higher α values in the Pareto environment allowed better discriminability between the more frequently encountered lottery options, also indicated by our model-free analysis (Table 1), which was crucial given the correlational structure between the two valuations. However, this result was not fully replicated in the pooled estimates: when estimating each environment separately, we found that recovered parameters were almost identical ($\alpha_{uniform}$=0.93, α_{Pareto}=0.92, Table S4, second and third columns), but a full model with random effect for the Pareto environment (similarly to the one run on M), revealed that there was a tuning of the function curvature when switching between environments (Table S4, rightmost column, p<0.001).

The power utility model is not designed to capture the dependence of the subjective value function on the distribution of valuations and hence we did not anticipate an adaptation of the function's curvature. Indeed, when comparing estimates of r across the two environments, we obtain inconclusive results: while the pooled estimates indicated higher r values in the Pareto environment (Table S4), the subject-level estimates point in the opposite direction (Fig. 4C, one-sided Wilcoxon sign-rank test between subject-level estimates of r, Z=0.0511, p=0.4796).

To conclude this section, we found that subjects adapted the DN encoding function's parameters to the two environments in line with our hypothesis, showing context-dependency in choice.

[Insert Fig. 4 here]

Discussion

In this study, we tested how the distributional properties of choice environments affect value encoding. In particular, we were interested in whether the subjective value of rewards is encoded via a mechanism such as divisive normalization (DN) exclusively in the Pareto environments for which it is efficient, or whether a DN representation is also employed in environments characterized by different reward distributions. To this end, we designed an experiment in which subjects were asked to make choices in two distinct statistical environments. In one environment, rewards were drawn from a Pareto distribution of valuations, for which DN is considered an efficient encoder. In the other environment, valuations were uniformly distributed and DN would not represent the environment most efficiently (*26*).

Our results indicate that subjects in our study were better described as using a DN mechanism than a power utility mechanism to encode the subjective value of rewards, no matter from which of our two distributions the rewards were drawn. As expected, the key parameter of the model tracked the median of the distribution. A model-free analysis indicated that, when in the Pareto environment and compared with the uniform environment, subjects made fewer mistakes in choice sets drawn from the center of the distribution at the expense of the margins. This is a principal property of the DN function. We then fitted our subjects' choices with two classic stochastic choice models – one was a standard RUM with a power utility function, and the other was a RUM with a utility function belonging to the family of DN models. Our subject-level and pooled model-fitting results suggested that the DN model better captured subjects' choice patterns in both the Pareto and the uniform environments (Table 2, Fig. 4C-D and Table S5). In line with the actual statistical properties of the two environments, subjects had higher reward expectations in the uniform environment. Taken together, these findings suggest that although subjects' choices were affected by the context of the choice environment, their choice mechanisms were constrained to a DN encoding of value (Fig. 5).

[Insert Fig. 5 here]

Our findings indicate that in both of our environments the encoding of value was better described by a DN than by a power utility encoder. One possible explanation is that Pareto distributions are common in the real world, and hence the brain has evolved a constraint that accords well with natural environments. Indeed, numerous sensory stimuli are characterized by Pareto-like statistical properties (*1*, *14*, *18*, *27*). On a larger scale, Pareto distributions also capture various ecological quantities, such as temporal and spatial measures of biodiversity (*37–40*). This is true also for environments that are related to value-based decisions, because many economic and financial quantities in modern societies (*41*, *42*), including consumption of several categories of consumer goods (*43*), have Pareto-like properties. An alternative explanation is that even the uniform distributions we examined are more efficiently dealt with by a DN encoder than by the non-divisive power utility encoder. Indeed, recent theoretical advances (*13*) have shown that all encoders with limited precision or accuracy must incorporate an implicit divisive cost, a class of encoders of which DN is a member.

Another important finding is that, compared with the standard utility functions used in economics, DN provides the brain with a highly flexible tool for the representation of choice options (*32*, *33*). Given the specific parameterization we employed for DN, our model embeds the standard concave utility function, but it is also suitable for capturing preferences that follow *S-shaped* functions, similar to the one suggested by Prospect Theory (*44*). DN further tracks the median of rewards (*expectations*), which allows for scale-invariant adjustments to different environments while ensuring a fine discrimination between stimuli that are in the center of the distribution (*2*, *9*, *13*, *45*). These adjustments – also evident in our data – give rise to context effects in choice processes (*25*, *33*, *46*, *47*) and are also the core reason for some notable perceptual illusions (*48*, *49*).

Our findings imply that some choice patterns should not be regarded as built-in decision biases, errors, or mistakes. Rather, they reflect adjustments of the brain, as a constrained system, to its environment, thus reflecting a rational value-encoding mechanism (*2*, *13*). Such an observation can explain the under-sampling of rare events when subjects adjust to new choice environments (*50*, *51*), since the main focus of the system is on the center of the distribution. On a broader view, our results could explain heterogeneity in individual decisions, such as the effect of one's position along the long-tail distributions of socioeconomic measures (and their shapes) on the quality of healthcare (*52*), savings (*53*), and consumption (*54*) choices.

Finally, an interesting question that stems directly from our research is to what extent our results generalize beyond decision making processes to other cognitive functions, such as sensory processing. Even though various natural sensory stimuli are described by Pareto-like properties (*14*, *18*, *45*), we also frequently encounter, and are required to process, non-natural non-Pareto stimuli (*55*, *56*). Our findings therefore invite further investigation into the effects of an obligate DN encoding on the sensory processing of non-Pareto stimuli.

Materials and Methods

Some of the data in this manuscript have been used in the conference paper in reference (*57*).

Experimental Design

Valuation task (STAGE I). Our goal was to establish whether the brain employs different value encoding models in environments with different reward distributions. To eliminate any additional prior heterogeneity in subjects' subjective valuations of money, we generated distributions of rewards in the subjective value (SV) space instead in dollar amounts (or expected values). To map the subject-specific SV space, we first recovered individual-specific subjective value functions over dollar amounts. To do this, in STAGE I, we used a valuation task, in which subjects reported their willingness to pay to participate in a lottery. See Table S1 for the list of 33 lotteries used in this task. On each trial, subjects were presented with a visualization of a 50-50 lottery on the computer screen and had to type in their willingness to pay to participate in it as a dollar amount (Fig. 2A). For each lottery, the valuation could range between the current lottery's minimal and maximal payoff, in \$0.10 increments. All subjects completed the same 33 trials in an order randomized at the subject level. At the end of the session, the realization of one randomly selected trial was implemented for payment, using a Becker–DeGroot–Marschak (BDM) (*58*) procedure which was designed to elicit truthful valuations.

Choice task (STAGE II). STAGE II was designed to test whether the distribution of rewards (lotteries with different subjective valuations) in a choice environment affects what value encoding model subjects use. Subjects were asked to choose the 50-50 lottery they preferred from two available options that varied from trial to trial. Lottery payoffs ranged between \$0 and \$60 in \$0.10 increments. Overall, subjects made 640 binary choices that were divided into two blocks of 320 trials each and presented on subsequent days. Our experimental manipulation was that in each block, the valuations were drawn either from a Pareto Type III distribution for which DN is an efficient code (*26*) or from a uniform distribution (Fig. 2A-B). The order in which subjects experienced these environments was counter-balanced across subjects. One trial was randomly selected for payment at the end of each experimental session. Subjects also completed additional 640 trials with six-option choice sets with lottery valuations drawn either from a Pareto Type III or uniform distributions. Thus, in total, in each environment subjects encountered two 320 choice blocks. The six-option blocks were designed to examine another research question that is beyond the scope of the current study and will be reported in a separate paper. Blocks were presented in an order randomized across subjects but on a given day, all blocks were drawn from the same distribution. Payments for STAGE II included a realization of one choice from each of the two sessions, and could be drawn either from the two-options sets or from the six-options sets.

Subjective Value of Money. We used each subject's STAGE I single lottery valuations to estimate their subjective value function over money. We expressed each subject i's subjective value of a 50-50 lottery that paid y_1 or y_2, each equally likely, using a power utility function as:

(iii) $$u_i(y_1, y_2) = 0.5y_1^{\rho_i} + 0.5y_2^{\rho_i}$$

If the curvature parameter $\rho_i < 1$, then subject i is risk-averse. When $\rho_i = 1$, the subject is risk-neutral. If $\rho_i > 1$, the subject is risk-seeking. Therefore, the certainty equivalents (c) that participants stated were converted to subjective values using the same power utility function such that $c = u^{1/\rho}$. We ran an NLS regression to estimate the ρ parameter separately for each subject.

We used the subject's estimated ρ_i, to pick different combinations of lottery dollar payoffs to create lotteries that had a specific SV to that individual. This enabled us to generate sets of lotteries whose implied SV distributions matched our target distributions (see below), regardless of individual differences in the curvature of the subjective value function.

Distributions of Valuations

Uniform Distributions of SVs. For each subject i, we computed the upper bound of the distribution as the SV of the maximal possible monetary payoff in the study, which was \$60 (i.e., $u_i^{max} = 60^{\rho_i}$). We then divided the range $[0, u_i^{max}]$ into 40 equally-spaced SV increments. For each of the increments, we created eight different lotteries, which would give the subject the subjective value in exactly this bracket (for a total of 320 lotteries). Since the joint distribution of a two-dimensional uniform distribution is independent, and hence determined by its marginals, we then picked pairs of lotteries from this set for generating binary choice sets.

Pareto Type III Distributions of SVs. The DN encoding function is information-maximizing for a bivariate Pareto distribution with a joint pdf $f_{u_i}(u_{i,1}, u_{i,2})$ (see Eq. 7 with $\mu_1 = 0$ in reference (*26*) to match the lower bound of the uniform distribution, and to avoid negative valuations). For every subject i and $k \in \{1,2\}$ is an index indicating the choice option within the choice set.

(iv) $$f_{u_i}(u_{i,1}, u_{i,2}) = \beta^2 \frac{2\left(\prod_{k=1}^{2} \frac{1}{\sigma_{i,1}} \left(\frac{u_{i,k}}{\sigma_{i,k}}\right)^{\beta-1} \right)}{\left(1 + \sum_{k=1}^{2} \left(\frac{u_{i,k}}{\sigma_k}\right)^{\beta}\right)^3},$$

and the marginal pdf being a univariate Pareto Type III pdf:

(v) $$f_{u_{i,k}} = \beta \frac{\frac{1}{\sigma_{i,k}} \left(\frac{u_{i,k}}{\sigma_{i,k}}\right)^{\beta-1}}{\left(1 + \left(\frac{u_{i,k}}{\sigma_{i,k}}\right)^{\beta}\right)^2},$$

(vi) $$E(u_{i,k} | u_{i,l}) = \sigma_{i,k} \left[1 + \left(\frac{u_{i,l}}{\sigma_{i,l}}\right)^{\beta}\right]^{1/\beta} \frac{\Gamma(2 - \frac{1}{\beta})\Gamma(\frac{\beta+1}{\beta})}{\Gamma(2)}, \forall k \neq l$$

We matched, for each subject, $E(u_{i,k} | u_{i,l})$ to the expectation of the uniform distribution, which was \$30 (and $\bar{u}_i = 30^{\rho}$ in SV-space), where Γ indicates the gamma function.

Following Proposition 4 in(*26*) and using the subject-specific parameterization, we generated the Pareto Type III distributions as a scale mixture of transformed exponential (or Weibull) random variables, so that:

(vii) $$u_{i,k} = \sigma_{i,k} \left(\frac{U_{i,k}}{Z_i}\right)^{\frac{1}{\beta}}, \quad for\ k \in \{1,2\}$$

where $U_k \sim Exp(\lambda = 1)$ and $Z \sim Exp(\lambda = 1)$ independently of all U_k. Fig. 2C presents three examples for such distributions with different ρ_i values.

Note that using only 320 draws may lead to under-sampling of the distributions. Therefore, to fully capture the shape of the distribution, for each subject, we first generated joint Pareto distributions with 100K draws. We then created small 600-draw experimental distributions that matched the large 100k-draw distributions, allowing a deviation of up to 0.2 utils from the actual first and second moments (mean and standard deviation) of the large 100k-draws sets. Fig. 2D compares matched and unmatched small sets, corresponding to the large 100k-draws set presented in Fig. 2C (middle panel). Finally, we truncated the long tail of the Pareto Type III distributions at $u_i^{max} = 60^{\rho}$ (eliminating 6.5 to 23.83 percent of the distribution, depending on the ρ parameter, the curvature of the subjective value function), to match the upper bound of the uniform distribution and to avoid extreme reward amounts. We then casted 320 SVs at random from the remaining valuations, which constituted the experimental subject-specific Pareto distributions.

Generating Binary Choice Sets from the Distributions of Valuations. The final step was to generate lottery dollar amounts from the SV distributions. For each lottery k with a valuation (u_k), we first randomly drew the first monetary payoff $x_{1,k}$ from a range of possible payoffs $\$0 - x^{max}$ in \$0.10 increments. We had to restrict the maximum value of $x_{1,k}$ to make sure that including it in the lottery, does not exceed the lottery valuation (u_k), and thus to avoid negative values for the second lottery payoff. We determined the maximal value of the first payoff $x_{1,k}$ using the minimum function:

$$\text{(viii)}\quad x_{1,k}^{max} = \min\{(2u_k)^{\frac{1}{\rho}}, 60\}.$$

We then solved for $x_{k,2}$, giving rise to the desired u_k, rounded to one decimal place, using the following equation:

$$\text{(ix)}\quad x_{2,k} = \left(2u_k - (\,x_{1,k})^{\rho}\right)^{\frac{1}{\rho}}.$$

Fig. S1 shows how the heterogeneity in ρ values affected the distributions of $x_{k,1}$ and $x_{k,2}$.

We restricted the share of trials with first-order stochastic dominance (FOSD) (trials on which both lottery payoffs of one lottery were higher or equal to the other lottery's payoffs) to 45 percent. For subjects with $\rho_i \to 0$, we could not generate experimental sets with only 45 percent of the trials. Thus, we fixed ρ_i=1, for all subjects with $\rho_i < 0.1$ (a total of 4 subjects, see Table S2), limiting the interoperability of data from this small number of subjects. In contrast, for two subjects with very high ρ's ($\rho_i > 4$), we also had to fix $\rho_i = 1$ in STAGE II of the study, since a very large tail from their Pareto distribution of SVs exceeded $60. Respectively, the interoperability of data from this subject is also limited. Nonetheless, we wanted to avoid any unjustified elimination of data, and therefore analyzed data from these six subjects. Importantly, our main qualitative findings do not change once we remove these subjects from our sample.

Procedures

Sessions. Experimental sessions were carried out online via Zoom while subjects completed the task on a website. We ran eight sessions of the experiment between May 2022 and August 2022. After instruction, subjects had to successfully answer a set of comprehension questions about the procedure before starting STAGE I. They could participate in STAGE II of the study only if they completed all trials in STAGE I. Subjects received all payments after completing both STAGE I and STAGE II. Subjects received a $10 participation fee and on average $24.5 in STAGE I (range $0-60) and $76.02 in STAGE II (range $7.3-120) from the decision task. All amounts are in Australian dollars. All parts of the experiment were self-paced. Both the valuation and the choice tasks were programmed in the oTree software package (*59*).

Participants. We recruited participants from various departments at the University of Sydney. Subjects gave informed written consent before participating in the study, which was approved by the local ethics committee at the University of Sydney. Seventy-six subjects (44 females, mean age=21.8, std: 3.34, range: 18-30) passed the comprehension questions and completed STAGE I and the two choice tasks of STAGE II.

Model Fitting

Sample-level (pooled) estimates. We estimated subjects' aggregated choice data via a probit choice function with maximum likelihood estimation (MLE). Standard errors were clustered at the subject level. Thus, in the pooled estimation subjects were treated as one representative decision-maker. In this analysis we used lotteries' monetary rewards (as opposed to their subjective valuations) to allow meaningful estimates of DN's M parameter, and to confine the range of lottery payoffs. For both DN and power utility, we report the results from models estimated on the full dataset and separately on each choice environment. To test the possibility of adaptation of the encoding function to the choice environments, we further report the results from three additional models estimated on the full dataset, which also included a dummy variable indicating the Pareto environment for the reward expectation, M parameter (DN) as $M = constant + M_{Pareto}xPareto$ and similarly for the functions' curvature parameters α (DN) and r (power utility), respectively.

Subject-level estimates. *DN.* In each choice environment, we recovered subject-specific estimates of the free parameters, restricting the search space as follows: $\alpha \in [0.1{,}5]$, $M \in [0, u_i^{max}]$ and $\theta > 0$ (see equation (i) in the text). We employed MLE using the Nelder-Mead algorithm with a max-iteration limit of 1,000 and a stopping criteria of 0.5 tolerance. We initialized M to the distributions' medians. θ was initialized at 0.03, matching the sample-level pooled estimate (see Table S5). For the α parameter, we took ten random initializations in the range {0.1,5} with a precision of 5. For calculating the likelihoods, in each of the 320 trials, we generated 10,000 samples with randomly drawn Gaussian noise. The log-likelihood function was thus given by –

$$\text{(x)}\quad log\mathcal{L}(\alpha_i, M_i, \theta_i | u_{i,t}) = y_{i,t}\log\left(Pr(y_{i,t} = 1 | u_{i,t})\right) + (1 - y_{i,t})log(\Pr(y_{i,t} = 0 | u_{i,t})),$$

where $y_{i,t} = \{0,1\}$ indicates the subject's i choice in trial $t = \{1, \ldots, 320\}$.

Power utility. We fitted the power utility model to recover subject-specific estimates of the r and θ parameters using a similar procedure. We restricted the search space as follows: $r \in \{0.1{,}5\}$, and $\theta > 0$ (see equation (ii) in the text). θ was initialized at 0.03, matching the sample-level pooled estimate (see Table S5). For the r parameter, we took ten random initializations in the range {0.1,5} with a precision of 5. All other procedures were identical to the DN model.

Acknowledgements
This work was funded with a grant from the Australian Research Council (DP 190100489 to Tymula and Glimcher).

Authors Contributions
V.K.D., V.A., S.B., A.B., K.L., P.G. and A.T. designed the study. V.A. collected the data. V.K.D, S.S. and V.A. analyzed the data. V.K.D., P.G. and A.T. wrote the manuscript.

Data Availability
All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. All the data and code used in this manuscript is available at https://github.com/VeredKurtz/ESVT.

Competing Interests
The authors declare that they have no competing interests.

References

1. E. P. Simoncelli, B. A. Olshausen, Natural Image Statistics and Neural Representation. *Annu. Rev. Neurosci.* **24**, 1193–1216 (2001).

2. P. W. Glimcher, Efficiently irrational: deciphering the riddle of human choice. *Trends in Cognitive Sciences* **26**, 669–687 (2022).

3. C. Frydman, L. J. Jin, Efficient Coding and Risky Choice. *The Quarterly Journal of Economics* **137**, 161–213 (2022).

4. N. Stewart, S. Reimers, A. J. L. Harris, On the Origin of Utility, Weighting, and Discounting Functions: How They Get Their Shapes and How to Change Their Shapes. *Management Science* **61**, 687–705 (2015).

5. D. Alempaki, E. Canic, T. L. Mullett, W. J. Skylark, C. Starmer, N. Stewart, F. Tufano, Reexamining How Utility and Weighting Functions Get Their Shapes: A Quasi-Adversarial Collaboration Providing a New Interpretation. *Management Science* **65**, 4841–4862 (2019).

6. A. Tymula, H. Plassmann, Context-dependency in valuation. *Current Opinion in Neurobiology* **40**, 59–65 (2016).

7. M. W. Khaw, P. W. Glimcher, K. Louie, Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences* **114**, 12696–12701 (2017).

8. J. A. Heng, M. Woodford, R. Polania, Efficient sampling and noisy decisions. *eLife* **9**, e54962 (2020).

9. K. Louie, P. W. Glimcher, Efficient coding and the neural representation of value. *Annals of the New York Academy of Sciences* **1251**, 13–32 (2012).

10. R. Polanía, M. Woodford, C. C. Ruff, Efficient coding of subjective value. *Nat Neurosci* **22**, 134–142 (2019).

11. M. W. Khaw, Z. Li, M. Woodford, Cognitive Imprecision and Small-Stakes Risk Aversion. *The Review of Economic Studies* **88**, 1979–2013 (2021).

12. A. Rustichini, K. E. Conen, X. Cai, C. Padoa-Schioppa, Optimal coding and neuronal adaptation in economic decisions. *Nat Commun* **8**, 1208 (2017).

13. K. Steverson, A. Brandenburger, P. Glimcher, Choice-theoretic foundations of the divisive normalization model. *Journal of Economic Behavior & Organization* **164**, 148–165 (2019).

14. O. Schwartz, E. P. Simoncelli, Natural signal statistics and sensory gain control. *Nat Neurosci* **4**, 819–825 (2001).

15. M. Carandini, D. J. Heeger, Normalization as a canonical neural computation. *Nat Rev Neurosci* **13**, 51–62 (2012).

16. S. R. Olsen, V. Bhandawat, R. I. Wilson, Divisive Normalization in Olfactory Population Codes. *Neuron* **66**, 287–299 (2010).

17. D. J. Heeger, Normalization of cell responses in cat striate cortex. *Visual Neuroscience* **9**, 181–197 (1992).

18. O. Schwartz, E. Simoncelli, “Natural Sound Statistics and Divisive Normalization in the Auditory System” in *Advances in Neural Information Processing Systems* (MIT Press, 2000; https://proceedings.neurips.cc/paper/2000/hash/96c5c28becf18e71190460a9955aa4d8-Abstract.html)vol. 13.

19. M. Carandini, D. J. Heeger, J. A. Movshon, Linearity and Normalization in Simple Cells of the Macaque Primary Visual Cortex. *J. Neurosci.* **17**, 8621–8644 (1997).

20. D. A. Kaliukhovich, R. Vogels, Divisive Normalization Predicts Adaptation-Induced Response Changes in Macaque Inferior Temporal Cortex. *J. Neurosci.* **36**, 6116–6128 (2016).

21. C. T. Kello, G. D. A. Brown, R. Ferrer-i-Cancho, J. G. Holden, K. Linkenkaer-Hansen, T. Rhodes, G. C. Van Orden, Scaling laws in cognitive sciences. *Trends in Cognitive Sciences* **14**, 223–232 (2010).

22. N. Chater, G. D. A. Brown, Scale-invariance as a unifying psychological principle. *Cognition* **69**, B17–B24 (1999).

23. A. Rangel, J. A. Clithero, Value normalization in decision making: theory and evidence. *Current Opinion in Neurobiology* **22**, 970–981 (2012).

24. K. Louie, L. E. Grattan, P. W. Glimcher, Reward Value-Based Gain Control: Divisive Normalization in Parietal Cortex. *J. Neurosci.* **31**, 10627–10639 (2011).

25. K. Louie, M. W. Khaw, P. W. Glimcher, Normalization is a general neural mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences* **110**, 6139–6144 (2013).

26. S. F. Bucher, A. M. Brandenburger, Divisive normalization is an efficient code for multivariate Pareto-distributed environments. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2120581119 (2022).

27. M. J. Wainwright, O. Schwartz, E. P. Simoncelli, “Natural Image Statistics and Divisive Normalization” in *Probabilistic Models of the Brain*, R. P. N. Rao, B. A. Olshausen, M. S. Lewicki, Eds. (The MIT Press, 2002; https://direct.mit.edu/books/book/2733/chapter/73923/Natural-Image-Statistics-and-Divisive), pp. 203–222.

28. R. V. Solé, S. A. Levin, J. H. Brown, V. K. Gupta, B.-L. Li, B. T. Milne, C. Restrepo, G. B. West, The fractal nature of nature: power laws, ecological complexity and biodiversity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **357**, 619–626 (2002).

29. M. Newman, Power laws, Pareto distributions and Zipf’s law. *Contemporary Physics* **46**, 323–351 (2005).

30. D. Mcfadden, Economic Choices. *The American Economic Review* **91** (2001).

31. Jacob Marschak, Binary Choice Constraints on Random Utility Indicators (1959).

32. P. W. Glimcher, A. A. Tymula, Expected subjective value theory (ESVT): A representation of decision under risk and certainty. *Journal of Economic Behavior & Organization* **207**, 110–128 (2023).

33. R. Webb, P. W. Glimcher, K. Louie, The Normalization of Consumer Valuations: Context-Dependent Preferences from Neurobiological Constraints. *Management Science* **67**, 93–125 (2021).

34. E. T. Rolls, F. Grabenhorst, G. Deco, Choice, difficulty, and confidence in the brain. *NeuroImage* **53**, 694–706 (2010).

35. O. Schwartz, A. Hsu, P. Dayan, Space and time in visual context. *Nat Rev Neurosci* **8**, 522–535 (2007).

36. Peter Wakker, Explaining the characteristics of the power (CRRA) utility family. *Health Economics* **17**, 1329–1344 (2008).

37. Sophia I. Passy, Power law relationships among hierarchical taxonomic categories in algae reveal a new paradox of the plankton. *A Journal of Macroecology* **15**, 528–535 (2006).

38. B. J. Enquist, Universal scaling in tree and vascular plant allometry: toward a general quantitative theory linking plant form and function from cells to ecosystems. *Tree Physiology* **22**, 1045–1064 (2002).

39. B. J. McGill, R. S. Etienne, J. S. Gray, D. Alonso, M. J. Anderson, H. K. Benecha, M. Dornelas, B. J. Enquist, J. L. Green, F. He, A. H. Hurlbert, A. E. Magurran, P. A. Marquet, B. A. Maurer, A. Ostling, C. U. Soykan, K. I. Ugland, E. P. White, Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters* **10**, 995–1015 (2007).

40. Mark Q. Wilber, Justin Kitzes, John Harte, Scale collapse and the emergence of the power law species–area relationship. *A Journal of Macroecology* **24**, 883–895 (2015).

41. X. Gabaix, Power Laws in Economics and Finance. *Annual Review of Economics* **1**, 255–294 (2009).

42. X. Gabaix, Zipf's Law and the Growth of Cities. *American Economic Review* **89**, 129–132 (1999).

43. B. J. Kim, V. Singh, R. S. Winer, The Pareto rule for frequently purchased packaged goods: an empirical generalization. *Mark Lett* **28**, 491–507 (2017).

44. D. Kahneman, A. Tversky, "Prospect Theory: An Analysis of Decision Under Risk" in *Handbook of the Fundamentals of Financial Decision Making* (WORLD SCIENTIFIC, 2012; https://www.worldscientific.com/doi/abs/10.1142/9789814417358_0006)vol. Volume 4 of *World Scientific Handbook in Financial Economics Series*, pp. 99–127.

45. E. P. Simoncelli, B. A. Olshausen, Natural Image Statistics and Neural Representation. *Annual Review of Neuroscience* **24**, 1193–1216 (2001).

46. J. S. Trueblood, S. D. Brown, A. Heathcote, J. R. Busemeyer, Not Just for Consumers: Context Effects Are Fundamental to Decision Making. *Psychol Sci* **24**, 901–908 (2013).

47. A. Tversky, I. Simonson, Context-Dependent Preferences. *Management Science* **39**, 1179–1189 (1993).

48. J. H. Reynolds, D. J. Heeger, The Normalization Model of Attention. *Neuron* **61**, 168–185 (2009).

49. O. Schwartz, R. Coen-Cagli, Visual attention and flexible normalization pools. *Journal of Vision* **13**, 25 (2013).

50. R. Hertwig, G. Barron, E. U. Weber, I. Erev, Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychol Sci* **15**, 534–539 (2004).

51. R. Hertwig, I. Erev, The description–experience gap in risky choice. *Trends in Cognitive Sciences* **13**, 517–523 (2009).

52. B. Handel, J. Kolstad, T. Minten, J. Spinnewijn, The Socio-Economic Distribution of Choice Quality: Evidence from Health Insurance in the Netherlands.

53. A. K. Shah, S. Mullainathan, E. Shafir, Some Consequences of Having Too Little. *Science* **338**, 682–685 (2012).

54. A. K. Shah, E. Shafir, S. Mullainathan, Scarcity Frames Value. *Psychol Sci* **26**, 402–412 (2015).

55. V. J. Barranca, X. G. Zhu, A computational study of the role of spatial receptive field structure in processing natural and non-natural scenes. *Journal of Theoretical Biology* **454**, 268–277 (2018).

56. C. Heurteloup, A. Merchie, S. Roux, F. Bonnet-Brilhault, C. Escera, M. Gomot, Neural repetition suppression to vocal and non-vocal sounds. *Cortex* **148**, 1–13 (2022).

57. V. Kurtz-David, V. Alladi, S. Bucher, A. Brandenburger, K. Louie, P. Glimcher, A. Tymula, "Choosers Adapt Value Coding to the Environment, But Do Not Attain Efficiency" in *Proceedings of the Annual Meeting of the Cognitive Science Society* (2023; https://escholarship.org/uc/item/7cp9r5hc)vol. 45.

58. G. M. Becker, M. H. DeGroot, J. Marschak, Measuring utility by a single-response sequential method. *Syst. Res.* **9**, 226–232 (1964).

59. D. L. Chen, M. Schonger, C. Wickens, oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* **9**, 88–97 (2016).

Figures and Tables

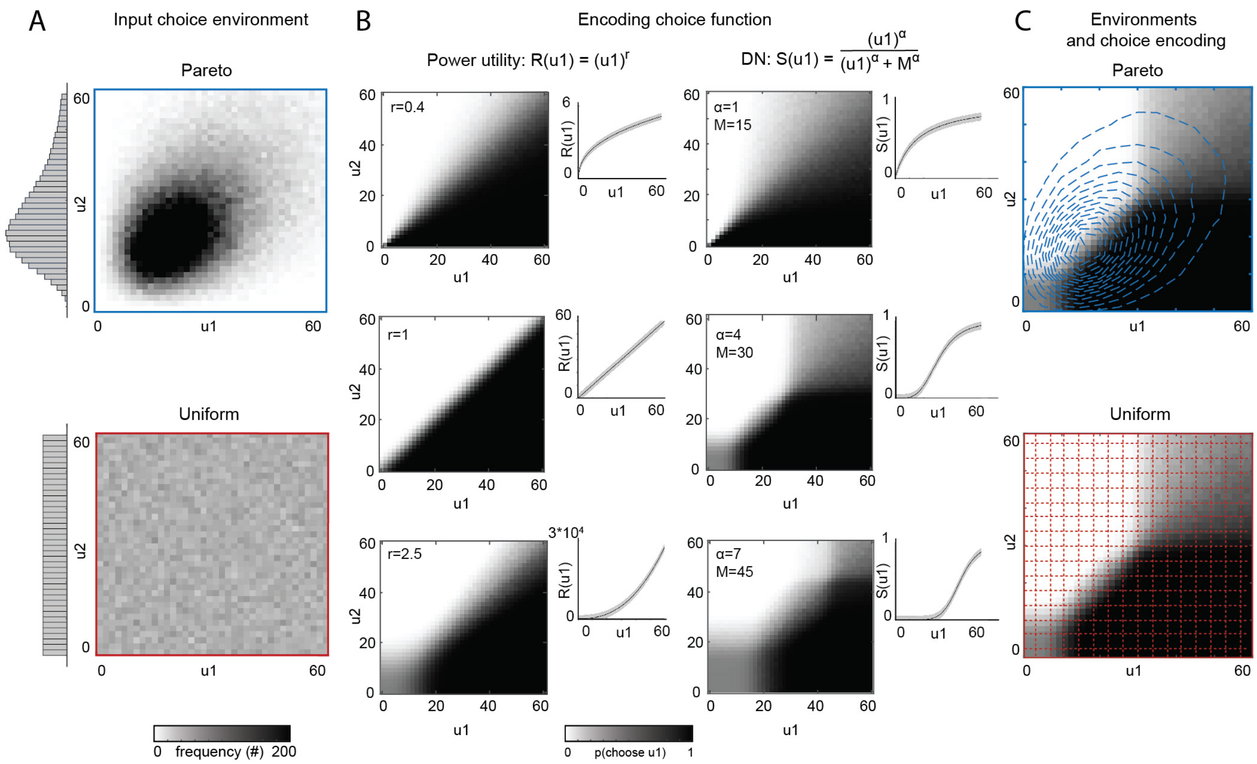

Figure 1. Research Question. (A) Choice environments are determined by the distribution of valuations. We compare a long-tailed bivariate Pareto Type III environment for which DN is efficient code with a uniformly distributed environment for which DN is not an efficient code. Figures show 2D histograms of simulated choice trials with valuations in the range $u_k \in [0{,}60]$ for every lottery $k \in \{1{,}2\}$. Each reward's value was drawn from 40 bins. Insets show their corresponding marginal distributions. We simulate 100k valuations per environment. See Materials and Methods for further details. (B) Value encoding choice functions. We test two different RUM models: classic power utility (left) and DN (right). The figure shows the probability of choosing a lottery with valuation u_1 over a lottery with valuation u_2 for various parameter values in each model. Insets show the subjective representation of u_1 in power utility (R), and in DN (S). For every combination of u_1 and u_2, we simulate 1k binary choice sets. We allow stochasticity in choice by incorporating additive noise, drawn from $\eta \sim N(0, 0.05 * \bar{R}^{max}(u_1)$, such that $\bar{R}^{max}(u_1)$ denotes the maximal subjective value of u_1 in the power utility model (and $\zeta \sim N(0, 0.05 * \bar{S}^{max}(u_1)$ in the DN model, respectively). We cast 10K noisy draws per simulated trial and reported average choice probabilities across simulated sets. (C) Contour plots indicate the mass of occurrences of (u_1, u_2) choice trial combinations in each environment. Contours were laid over a representative DN model with $\alpha = 4, M = 30$ (middle right panel in (B).

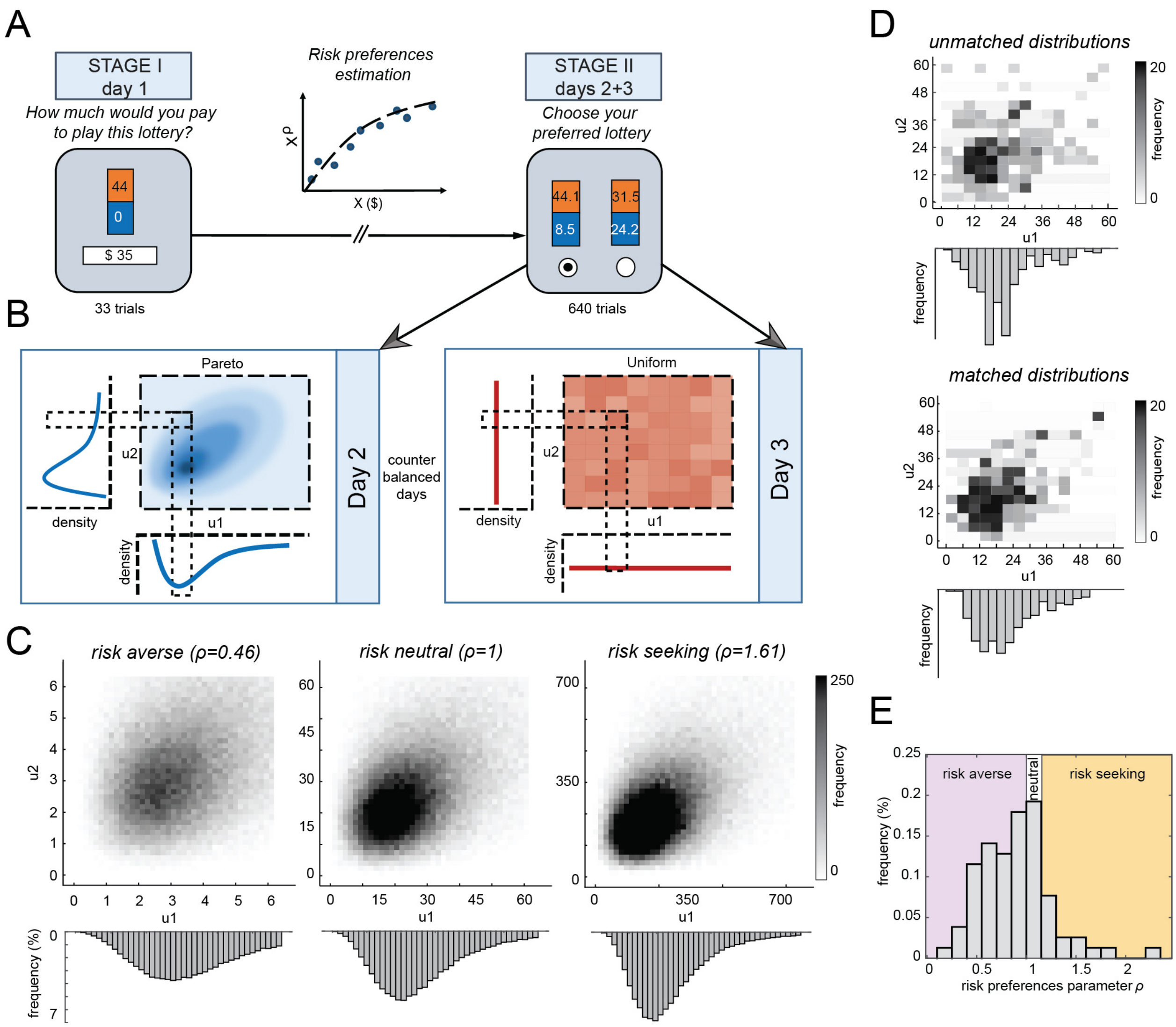

Figure 2. Experimental Design. (A) Timeline. In STAGE I, subjects reported their valuations for 33 lotteries. Valuations were used to recover the curvature of the subjective value function for each subject using NLS estimation. Based on those estimates, we generated subject-specific bi-dimensional uniform and Pareto Type III distributions of valuations for STAGE II of the study. In STAGE II, subjects completed two sets of 320 binary choices between 50-50 lotteries (640 choices in total). (B) Bi-dimensional Pareto and uniform distributions. In the uniform distribution, we created 40 bins of subjective values between 0 and the maximal payoff in the study (\$60, $u_i^{max} = 60^{\rho_i}$) with eight lotteries in each bin. We then picked pairs of lotteries from this set to create binary choice sets. In the Pareto distribution, we used a Gamma-weighted scale mixture of exponential random variables to capture the covariance structure of the bi-variate Pareto distribution. (C) Choice sets in STAGE II controlled for differences in individual subjective value function (risk attitudes), modulating the second moment (std) of the Pareto distribution (see eq. (vi) in Materials and Methods). The histograms show the bi-dimensional Pareto distributions and their marginals (with 100k draws per distribution) from three representative subjects: left - a risk averse subject, middle - a risk neutral subject, right - a risk seeking subject. (D) Experimental sets with 320 trials were prone to under-sampling (see top, unmatched distribution). We matched experimental sets to the distributional shape of a larger set with 100k draws (see bottom, matched distributions). The figure shows an example corresponding to the middle panel in (C). (E) Recovered estimates of subjective value curvature (risk attitudes) from STAGE I. See Materials and Methods and Fig. S1 for further details. See Table S2 for a list of the estimated subjective value function curvatures (risk parameter ρ).

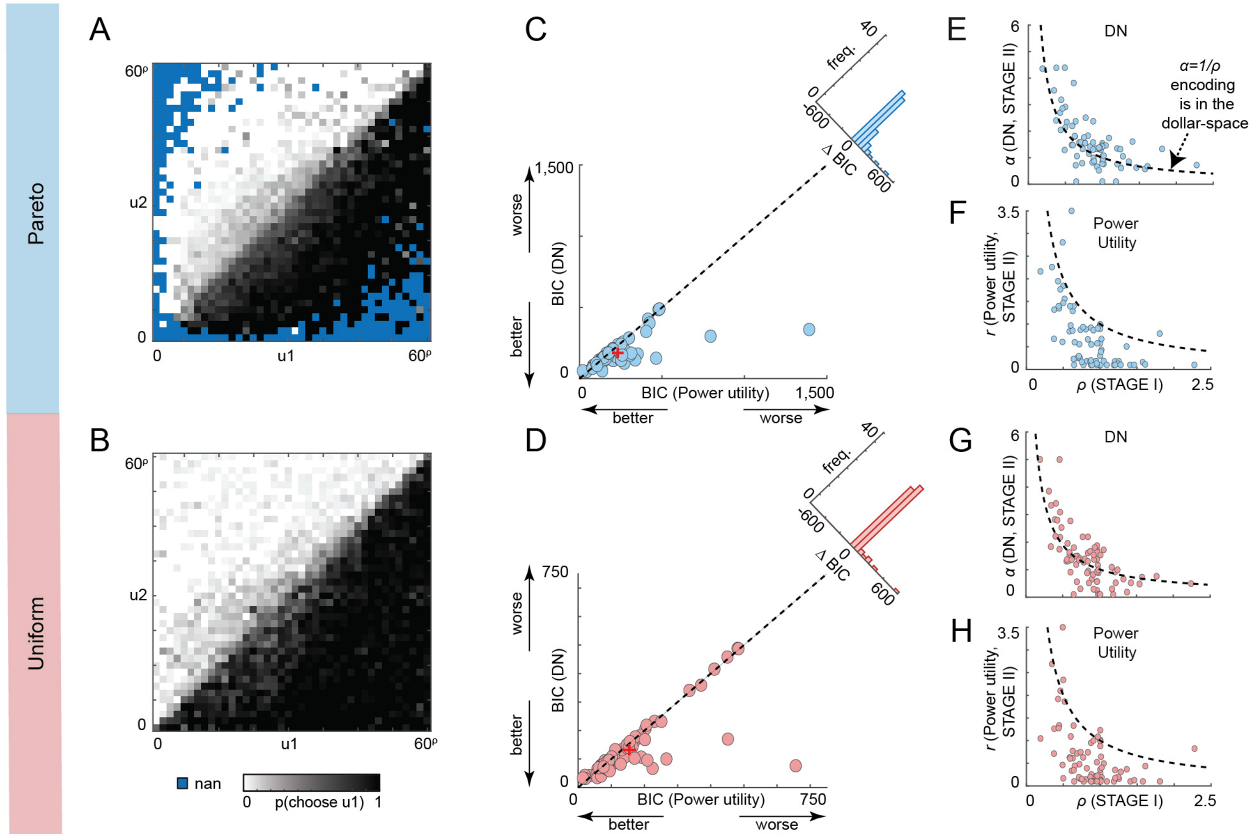

Figure 3. Model-fitting. (A-B) Probability of choosing lottery 1 with valuation u_1, in a (u_1, u_2) choice set. Data is aggregated over subjects. Within subjects, valuations are divided into 60 equally-spaced bins. (A) The Pareto environment. (B) The uniform environment. (C-D) Each dot is one subject's DN model BIC score (y-axis) plotted against the same subject's power utility BIC score (x-axis). A dashed 45-degree line indicates when both models are equally successful. Inset shows the difference in BIC scores ($BIC_{power} - BIC_{DN}$). (C) The Pareto environment. (D) The uniform environment. (E-F) Relationship between the STAGE I curvature of the subjective value function (ρ) and STAGE II subjective value functions in the Pareto environment. Dashed curve indicates hyperbolic function $y = \frac{1}{x}$. (E) DN model (α parameter). (F) Power utility model (r parameter). (G-H). Same as (E-F), but for the uniform environment. Dots indicate individual subjects, + indicate the sample averages. N=76.

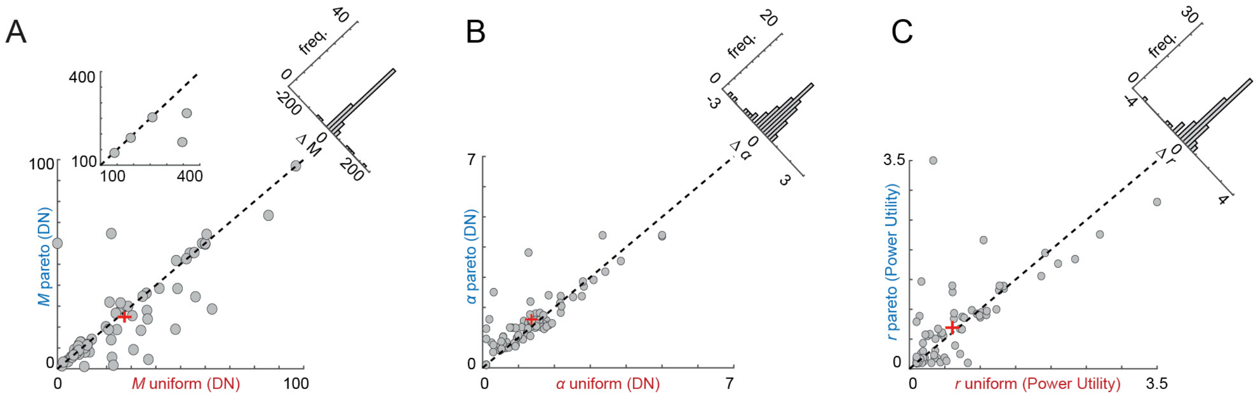

Figure 4. Cross-environment adaptation. (A-B) Adaptation of the encoding function in the DN model. (A) Best-fitting M parameter in the uniform (x-axis) vs. the Pareto (y-axis) environments. Estimates of M's are in utility space. Left inset: outliers. Right (diagonal) inset: Difference in the estimates of M across choice environments ($M(uniform) - M(Pareto)$). Insets do not show three additional (risk-seeking) subjects whose M's are >400 (in util units). Dots indicate individual subjects, + indicate sample average without the inset outliers, N=76. (B) same as (A) for the DN's α parameter. (C) Adaptation of the encoding function in the power utility model. Same as (B), but for the r parameter from the power utility model. (B-C) Dots indicate individual subjects, + indicate sample average, N=76.

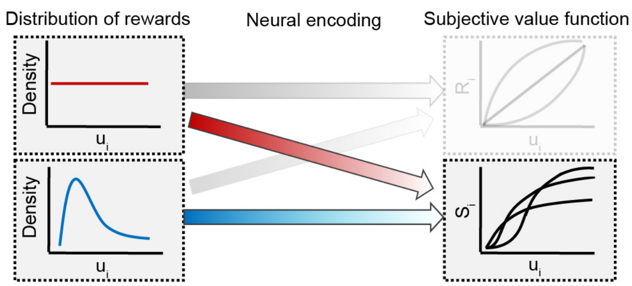

Figure 5. Summary of main findings.

Table 1. Results from the model-free analysis. Probit regressions with the dependent variable is equal to 1 when the subject chose the lottery with the higher SV, and zero otherwise. Column (1) model was run on the full sample. The independent variables are the absolute SV difference between the two lotteries, a dummy indicating the Pareto environment and their interaction. Column (2) model was run on data including choice sets in the center of the distributions. The model includes the same independent variables as model (1), and an additional dummy equal to 1 if the lottery was taken from around the diagonal (and zero otherwise, see text for definitions) and its interaction with the Pareto dummy. Standard errors clustered on subject in parentheses, + p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

	(1) Full sample	(2) Center of the distributions
SV difference	0.0002*	-0.0001
	(0.0001)	(0.0000)
Pareto	-0.3311***	-0.1479***
	(0.0405)	(0.0374)
Pareto*SV difference	-0.0001*	-0.0003***
	(0.0001)	(0.0001)
Near diagonal		-0.7400***
		(0.0692)
Pareto*Near diagonal		0.1742**
		(0.0597)
Constant	1.3260***	1.1201***
	(0.0704)	(0.0713)
N	48640	22442
pseudo R-sq	0.015	0.036

Table 2. Median estimates.

	Power Utility model			DN model			
Parameter	r	θ_p	BIC	α	M	θ_{DN}	BIC
Uniform	0.379	0.054	121.120	1.299	23.216	0.0233	104.670
Pareto	0.528	0.103	184.761	1.358	18.875	0.026	154.482

Supporting Materials

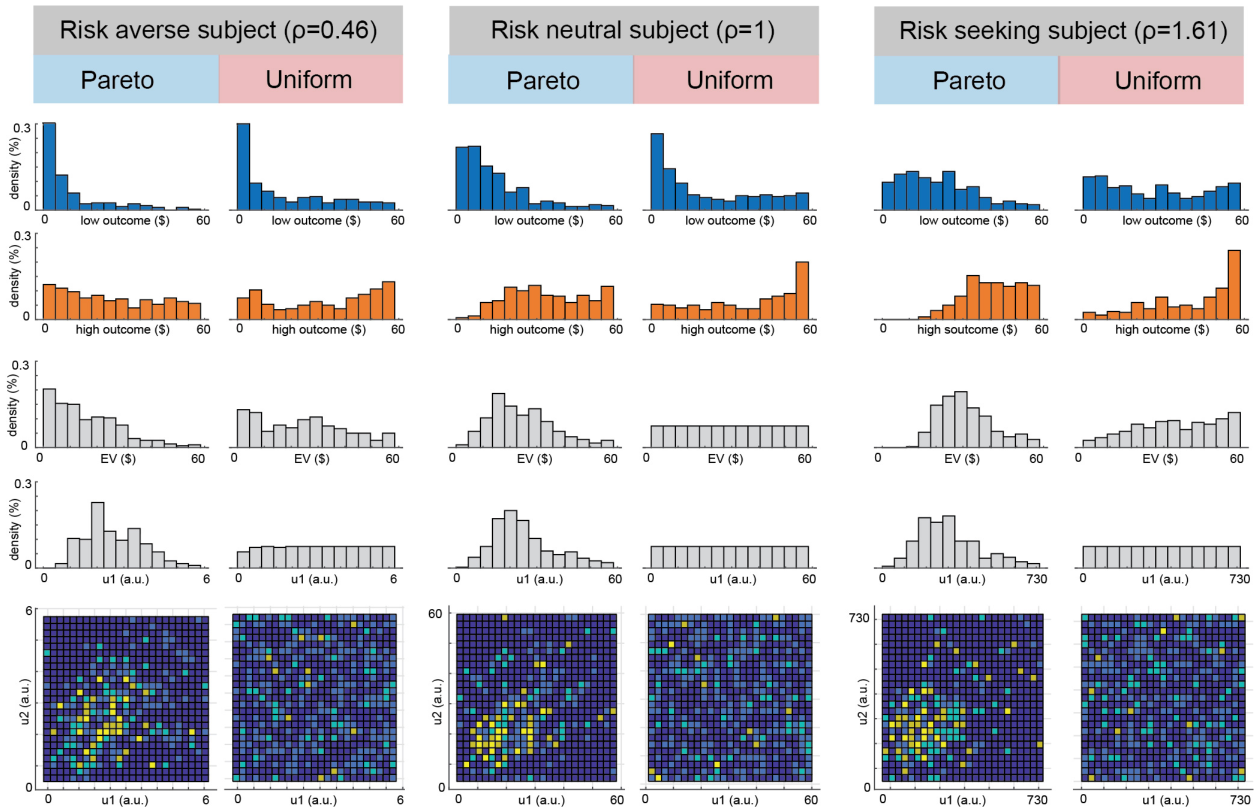

Figure S1. Representative Sets in STAGE II. Left – a risk averse subject, middle – a risk neutral subject, right – a risk seeking subject. Top to bottom: (1) Distributions of the high winning amount in Lottery 1 (in dollars); (2) Distributions of the low winning amount in lottery 1 (in dollars); (3) Distribution of the expected earnings (EV) of Lottery 1 (in dollars); (4) Distributions of the valuations (u1) of Lottery 1 (in util units); (5) 2-dimensional histogram of the valuations of Lottery 1 and Lottery 2 (u1 and u2, in util units).

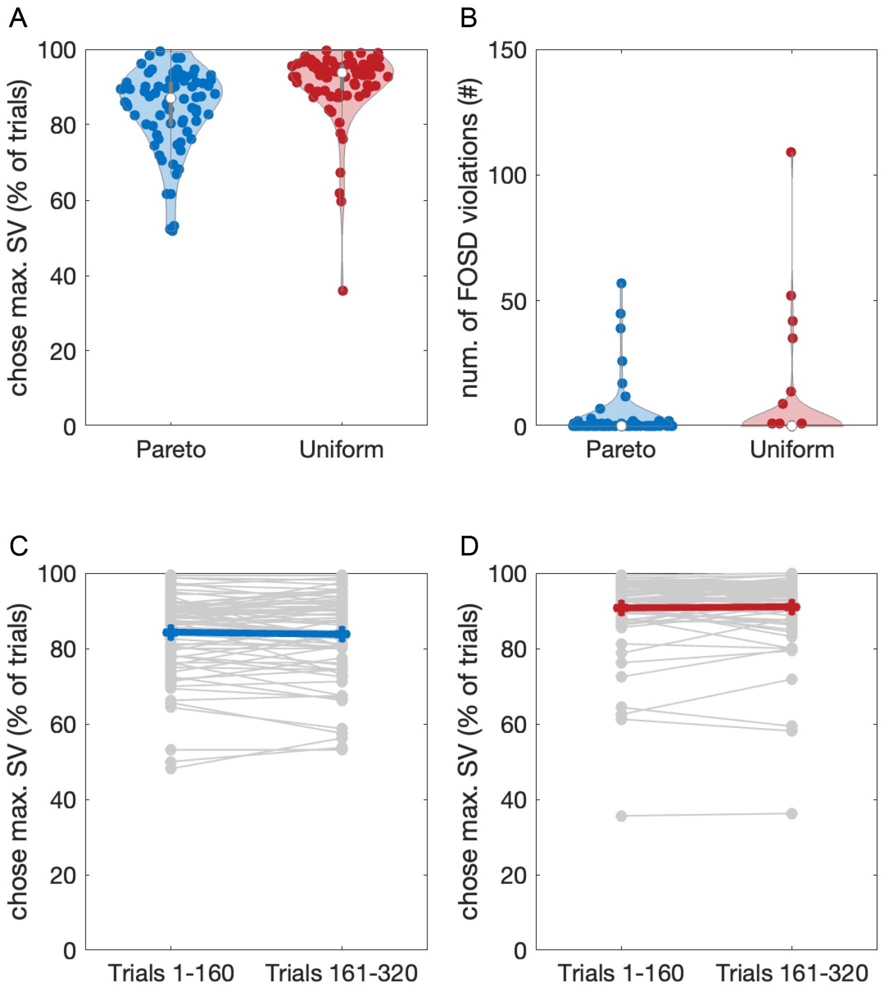

Figure S2. Descriptive statistics. (A) Violins show the share of trials in which subjects chose the lottery with the higher subjective value. (B) Violins show the number of FOSD violations per subject. Dots indicate individual subjects. N=76. (C-D) Share of trials in which subjects chose the lottery with the higher SV, first half of the session (trials 1-160), compared with the second half of the session (trials 161-320). Each gray line indicates a subject. Colored lines are sample averages. (C) Pareto distribution sessions. (D) Uniform distribution sessions.

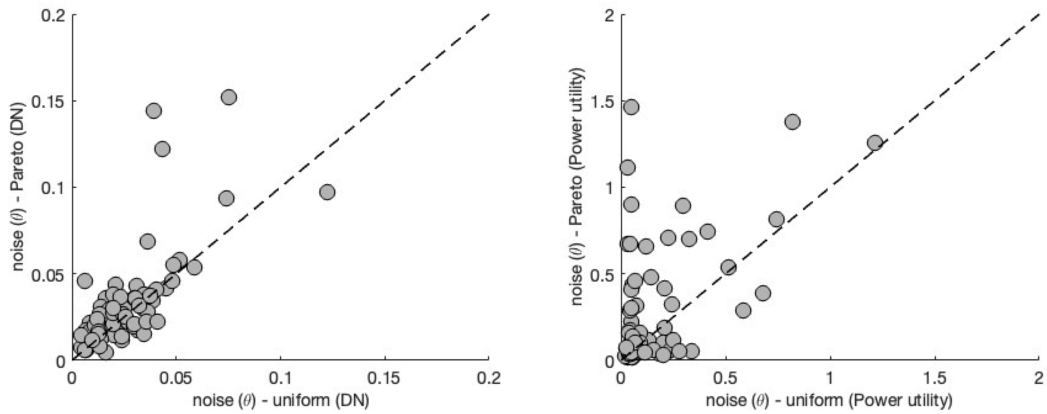

Figure S3. Noise estimates. Comparing the best-fitting σ parameter (decision noise) across the distributional environments reveals noise levels were higher in the Pareto environment. *Left* - DN model (one-sided Wilcoxon sign-rank test, Z=2.2314, p=0.0257). *Right* - Powe Utility model (one-sided Wilcoxon sign-rank test, Z=2.9172, p=0.0035). Scatters indicate individual subjects. N=76.

Table S1. Lotteries used in STAGE I

lottery	x1	x2
1	60	0
2	55	0
3	50	0
4	45	0
5	40	0
6	35	0
7	30	0
8	25	0
9	20	0
10	15	0
11	10	0
12	5	0
13	60	5
14	55	5
15	50	5
16	45	5
17	40	5
18	35	5
19	30	5
20	25	5
21	20	5
22	15	5
23	10	5
24	60	10
25	55	10
26	50	10
27	45	10
28	40	10
29	35	10
30	30	10
31	25	10
32	20	10
33	15	10

Table S2. Individual-level estimates of risk preferences from subjects' bids in STAGE I.

SID	ρ	SE	SID	ρ	SE
1005	0.000 (*)	0.009	1501	0.866	0.051
1104	0.000 (*)	0	1406	0.875	0.018
1205	0.000 (*)	0	1308	0.876	0.093
1401	0.000(*)	0	1614	0.876	0.068
1001	0.347	0.055	908	0.88	0.045
1006	0.374	0.067	1508	0.891	0.066
1301	0.377	0.056	1408	0.909	0.033
1201	0.429	0.106	909	0.927	0.176
1617	0.437	0.04	1008	0.946	0.055
903	0.443	0.045	1502	0.954	0.065
1619	0.461	0.036	1607	0.954	0.042
1605	0.463	0.094	1402	0.961	0.067
1601	0.49	0.057	1604	0.968	0.018
1101	0.492	0.069	1105	0.981	0.018
910	0.499	0.053	1204	1	0
1610	0.5	0.043	1603	1	0
1405	0.518	0.048	1409	1.008	0.132
906	0.586	0.039	1513	1.012	0.058
1407	0.593	0.045	1012	1.018	0.028
1611	0.601	0.054	1613	1.018	0.016
1106	0.612	0.058	1505	1.047	0.042
1109	0.627	0.044	1608	1.09	0.157
1307	0.628	0.039	1305	1.185	0.126
1403	0.639	0.06	1004	1.192	0.145
1615	0.64	0.084	1002	1.201	0.068
904	0.647	0.048	1504	1.203	0.142
1510	0.655	0.074	1102	1.244	0.086
902	0.686	0.047	1003	1.253	0.062
1107	0.711	0.227	1609	1.278	0.095
1303	0.718	0.038	1304	1.358	0.197
1103	0.727	0.038	1503	1.407	0.109
1010	0.733	0.041	1302	1.577	0.143
1512	0.752	0.025	1011	1.61	0.157
1506	0.769	0.036	1616	1.808	0.24
1507	0.775	0.074	912	2.276	0.185
1306	0.806	0.055	1202	124.189 (*)	28.955
907	0.819	0.045	1203	4.432 (*)	0.261
1014	0.833	0.026			

(*) For these subjects we could not generate distributions of valuations for STAGE II that would adhere to our requirement to limit the number of trials with FOSD violations (when $\rho_i \longrightarrow 0$), or without having to censor a very large tail of the Pareto distribution (when $\rho_i > 4$). Instead, for these subjects we plugged-in $\rho_i = 1$ to generate the distributions for STAGE II.

Table S3. Robustness checks for the findings presented in Column (2) in Table 1. We vary the definitions for *center of the distributions* (center) and *around the diagonal* (diagonal). Column (1) corresponds to the regression presented in the Main Text.

	(1) Center: \$9-42 Diagonal: $0.9 < \frac{u_2}{u_1} < 1.1$	(2) Center: \$9-42 Diagonal: $0.95 < \frac{u_2}{u_1} < 1.05$	(3) Center: \$12-39 Diagonal: $0.9 < \frac{u_2}{u_1} < 1.1$	(4) Center: \$12-39 Diagonal: $0.95 < \frac{u_2}{u_1} < 1.05$
SV difference	-0.0001	-0.0000	-0.0002**	-0.0001**
	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Pareto	-0.0003***	-0.0003***	-0.0002***	-0.0002**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Pareto*SV difference	-0.1479***	-0.1547***	-0.1138**	-0.1228***
	(0.0374)	(0.0334)	(0.0394)	(0.0354)
Near diagonal	-0.7400***	-0.8916***	-0.6463***	-0.8166***
	(0.0692)	(0.0730)	(0.0708)	(0.0740)
Pareto*Near diagonal	0.1742**	0.3213***	0.1427*	0.3178***
	(0.0597)	(0.0660)	(0.0656)	(0.0734)
Constant	1.1201***	1.0618***	1.0146***	0.9580***
	(0.0713)	(0.0635)	(0.0726)	(0.0642)
	22442	22442	16576	16576
N	0.036	0.027	0.029	0.022
pseudo R-sq	-0.0001	-0.0000	-0.0002**	-0.0001**

Table S4. Individual-level best-fitting model parameters across environments (STAGE II).

SID	ρ	DN						Power Utility			
		Pareto			Uniform			Pareto		Uniform	
		M	α	θ	M	α	θ	r	θ	r	θ
1001	0.35	4.063	3.54	0.031	4.142	3.851	0.014	2.258	1.257	2.692	1.215
1006	0.37	1.442	2.4	0.058	1.808	1.838	0.051	0.674	0.192	0.714	0.207
1301	0.38	2.713	4.39	0.122	2.52	3.342	0.043	1.319	0.747	1.265	0.413
1201	0.43	5.709	2.7	0.025	5.517	2.82	0.02	1.947	1.376	1.918	0.819
903	0.44	3.642	3.18	0.043	3.622	3.417	0.031	1.347	0.539	1.336	0.517
1617	0.44	5.583	2.05	0.03	5.938	2.195	0.02	1.397	0.711	1.274	0.227
1619	0.46	6.615	2.86	0.012	6.615	2.795	0.01	1.767	0.456	2.1	0.067
1101	0.49	7.491	4.39	0.014	7.491	5	0.024	2.803	0.032	3.5	50.499
1601	0.49	3.221	1.29	0.026	2.299	1.713	0.024	0.608	0.142	0.463	0.054
910	0.5	7.726	2.91	0.012	7.726	3.091	0.024	1.843	0.067	2.341	0.028
1610	0.5	7.439	2.14	0.046	7.732	2.158	0.048	1.466	1.463	0.169	0.051
1405	0.52	8.329	2.35	0.007	8.337	2.52	0.005	1.558	0.046	1.862	0.035
906	0.59	9.181	1.48	0.038	6.865	1.268	0.035	0.84	0.418	0.628	0.21
1407	0.59	10.038	2.19	0.032	11.066	1.331	0.032	1.397	0.041	0.605	0.187
1611	0.6	10.916	1.47	0.021	8.483	1.932	0.03	0.971	0.391	1.092	0.678
1106	0.61	11.711	3.82	0.069	12.268	1.277	0.037	3.5	0.016	0.335	0.054
1109	0.63	7.008	1.82	0.025	8.8	1.596	0.02	0.856	0.324	0.774	0.242
1307	0.63	12.895	1.8	0.021	9.504	1.48	0.028	1.29	0.9	0.61	0.05
1403	0.64	10.037	2.37	0.008	9.728	2.764	0.013	1.392	0.818	1.331	0.744
1615	0.64	8.884	0.71	0.023	8.685	0.911	0.041	0.231	0.048	0.348	0.114
904(**)	0.65	14.149	0.1	1	14.149	0.1	1	0.1	3.84	0.1	1.174
1510	0.66	12.807	1.06	0.02	9.392	1.328	0.03	0.599	0.162	0.444	0.041
902	0.69	8.798	1.57	0.033	8.454	1.776	0.018	0.249	0.051	0.717	0.224
1107	0.71	3.503	1.59	0.042	4.333	1.461	0.045	0.294	0.111	0.193	0.053
1303	0.72	8.09	1.55	0.027	13.044	1.146	0.018	0.556	0.158	0.479	0.093
1010	0.73	20.129	1.33	0.036	19.703	1.466	0.021	0.874	0.89	0.747	0.296
1103	0.73	7.997	0.79	0.03	11.232	0.562	0.025	0.255	0.073	0.162	0.044
1512	0.75	11.392	1.76	0.054	12.119	1.32	0.059	0.165	0.05	0.154	0.054
1506	0.77	18.808	1.55	0.024	20.815	1.373	0.012	0.935	0.676	0.63	0.045
1507(**)	0.77	1.15	2.51	0.055	10.924	0.306	0.049	0.1	0.101	0.1	0.09
1306	0.81	27.078	1.83	0.021	24.896	2.176	0.02	0.886	0.437	0.143	0.054
907	0.82	11.66	1.15	0.036	8.719	1.231	0.016	0.181	0.035	0.391	0.07
1014	0.83	18.769	1.42	0.03	24.162	1.355	0.034	0.663	0.288	0.742	0.587
1501	0.87	34.674	0.82	0.023	34.617	0.883	0.027	0.183	0.049	0.196	0.049
908	0.88	27.922	0.99	0.021	36.342	0.95	0.018	0.55	0.227	0.255	0.047
1308	0.88	2.713	4.39	0.122	36.133	2.176	0.03	0.1	0.061	0.1	0.033
1406	0.88	29.879	1.57	0.016	28.679	1.578	0.008	0.923	0.674	0.952	0.03
1614	0.88	31.894	1.31	0.022	21.161	1.267	0.036	0.2	0.074	0.1	0.026
1508	0.89	1.747	0.87	0.041	22.703	0.518	0.041	0.1	0.04	0.1	0.04
1408	0.91	38.529	1.34	0.015	41.378	1.638	0.004	0.865	0.656	1.001	0.121
909	0.93	18.397	0.66	0.035	33.938	0.65	0.039	0.156	0.052	0.114	0.044
1008	0.95	13.914	1.33	0.044	21.842	0.822	0.021	0.183	0.054	0.219	0.05
1502	0.95	18.941	0.52	0.021	47.922	0.293	0.011	0.1	0.022	0.1	0.022
1607	0.95	31.53	1.38	0.038	25.643	1.5	0.038	0.133	0.055	0.523	0.278
1402	0.96	25.401	1.35	0.036	30.41	1.333	0.031	0.575	0.3	0.225	0.053
1604	0.97	52.639	1.37	0.006	52.397	1.779	0.006	0.906	0.042	1.087	0.043
1105	0.98	4.505	1.3	0.046	36.939	0.1	0.006	0.1	0.03	0.1	0.038

SID	ρ	DN						Power Utility			
		Pareto			Uniform			Pareto		Uniform	
		M	α	θ	M	α	θ	r	θ	r	θ
1005	1(*)	55.271	1.04	0.017	53.736	0.573	0.007	0.582	0.286	0.29	0.045
1013	1(*)	60	1.54	0.004	58.779	1.346	0.016	0.984	0.319	0.881	0.075
1104	1(*)	5.542	0.83	0.015	22.165	0.247	0.009	0.1	0.014	0.134	0.033
1202	1(*)	59.606	1.44	0.015	60	1.698	0.009	0.999	1.115	1.229	0.03
1203(**)	1(*)	60	0.1	1	60	0.1	1	0.1	5.389	0.1	42.832
1204	1	59.894	1.46	0.006	60	1.603	0.007	0.947	0.041	1.016	0.034
1205	1(*)	9.181	1.02	0.028	30.88	0.569	0.036	0.1	0.021	0.127	0.05
1401	1(*)	34.454	0.77	0.015	56.078	0.928	0.012	0.305	0.071	0.364	0.053
1603	1	55.57	1.42	0.015	55.573	1.85	0.035	0.896	1.05	0.991	2.048
1409	1.01	23.842	0.96	0.144	36.602	0.1	0.039	0.1	0.106	0.1	0.204
1513	1.01	28.602	0.94	0.017	62.8	1.125	0.013	0.391	0.117	0.645	0.247
1012	1.02	64.681	0.76	0.018	21.882	0.885	0.025	0.431	0.175	0.191	0.047
1613	1.02	64.294	1.08	0.015	60.575	0.993	0.013	0.672	0.412	0.404	0.048
1505	1.05	26.584	1.46	0.027	23.729	1.501	0.02	0.187	0.059	0.495	0.236
1608	1.09	73.339	1.74	0.093	85.689	1.691	0.074	0.1	0.101	0.1	0.066
1305(**)	1.18	97.004	0.1	0.152	96.82	0.1	0.075	0.1	0.705	0.1	0.326
1004	1.19	103.965	0.693	0.013	59.74	0.796	0.026	0.352	0.135	0.278	0.116
1002	1.2	38.408	1.35	0.022	48.763	1.09	0.012	0.179	0.051	0.353	0.083
1504	1.2	137.745	0.76	0.027	137.959	0.299	0.013	0.1	0.041	0.177	0.048
1102	1.24	51.763	1.32	0.038	48.245	0.921	0.02	0.124	0.047	0.2	0.044
1003	1.25	40.119	0.62	0.015	143.863	0.527	0.02	0.179	0.04	0.159	0.049
1609	1.28	186.451	1.05	0.026	187.422	0.995	0.026	0.1	0.063	0.311	0.154
1304	1.36	254.695	0.85	0.016	257.866	0.665	0.012	0.505	0.483	0.368	0.144
1503	1.41	23.511	1.61	0.029	136.216	0.186	0.018	0.1	0.043	0.1	0.066
1606	1.46	175.342	0.62	0.036	351.572	0.526	0.023	0.1	0.048	0.126	0.048
1302	1.58	266.753	0.57	0.017	359.864	0.481	0.031	0.16	0.048	0.1	0.047
1011	1.61	730.53	0.67	0.012	719.768	0.729	0.014	0.111	0.056	0.437	0.338
1616	1.81	1154.46	1.33	0.097	1639.48	0.761	0.122	0.79	0.032	0.1	0.204
912	2.28	1917.44	0.72	0.022	7816.41	0.494	0.008	0.1	0.048	0.825	0.03

(*) These subjects had either a STAGE I estimate of $\rho_i = 0$ or $\rho_i > 4$. For those subjects we could not generate distributions of valuations for STAGE II that would adhere to our requirement to limit the number of trials with FOSD violations (when $\rho_i \longrightarrow 0$), or without having to censor a very large tail of the Pareto distribution (when $\rho_i > 4$). Instead, for these subjects we plugged-in $\rho_i = 1$ to generate the distributions for STAGE II.

(**) Subjects who had >20 FOSD violations in at least one of the treatments.

Table S5. Pooled estimates, dollar space. In practice, to allow a better identification of the model parameters, we estimated the parameter τ, such that $\tau = M^{\alpha}$. We recovered M post-hoc by simply plugging-in τ and α into the equation. Standard errors in parentheses, + p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Model	Parameter		All data	Uniform	Pareto	All data (α)	All data (M)
DN	α	Constant	0.9660***	0.9277***	0.9184***	0.8247***	0.8535***
			(0.1321)	(0.1763)	(0.0955)	(0.0815)	(0.0840)
		α_{Pareto}				0.1415***	
						(0.0399)	
	τ	Constant	42.3389***	49.1992**	40.0310***	42.8550***	65.5239***
			(11.6202)	(15.3108)	(9.0071)	(10.0872)	(16.8167)
		τ_{Pareto}					-29.2557**
							(11.2962)
	$M = \tau^{\frac{1}{\alpha}}$	Constant	48.3049	66.6531	55.5609	95.2584	134.3315
		M_{Pareto}					-52.225174
	θ		0.0918**	0.0920*	0.0705***	0.0760***	0.0641***
			(0.0284)	(0.0405)	(0.0160)	(0.0181)	(0.0148)
	BIC		32921.34	14239.49	18263.75	32492.85	32513.34
	N		48,640	24,320	243,20	48,640	48,640
Power Utility	r	Constant	0.5519***	0.5749***	0.5569***	0.5106***	
			(0.0393)	(0.0469)	(0.0396)	(0.0375)	
		r_{Pareto}				0.0696***	
						(0.0169)	
	θ		1.4318***	1.8637**	1.2005***	1.3520***	
			(0.3598)	(0.5752)	(0.2695)	(0.3218)	
	BIC		33113.26	14258.23	18375.14	32662.57	
	N		48,640	24,320	243,20	48,640	