



Inconsistency in risk preferences: a psychophysical anomaly

Ivo Vlaev*

Centre for Health Policy, Imperial College London, London, UK

*Correspondence: i.vlaev@imperial.ac.uk

There is a fundamental problem with the conceptualization of the individual agent in economics and public policy. At the heart of both domains is the notion that people should be able to make stable tradeoffs between different goods and quantities. For example, people might have to choose between the benefit gained investing money in a pension fund or in joining a gym to improve their health; or they may need to trade-off risk and return when choosing a pension fund. Should policy makers help people make the right choices when faced with such difficult decisions? More pertinently, policy makers themselves frequently need to arbitrate between dissimilar options. For example, a minister may be forced to decide to fund hospitals rather than schools, and in doing so trade-off the health and education level of the population. Recent theoretical and empirical work suggests, however, that people cannot make stable tradeoffs, i.e., independent of other available choice options or the context. This inability seems to reflect a basic property of human cognition that applies right across psychology, from the basic psychophysics of sound perception right through to high-level cognitive processes in judgment and decision making.

This cognitive limitation has implications for economics and public policy where it raises important questions for the central methodologies used to measure and derive human preferences. Popular methods like functional measurement (Anderson and Zalinski, 1988) and conjoint analysis (Louviere, 1988; Green and Srinivasan, 1990) measure tradeoffs by asking respondents for attractiveness ratings of options (e.g., policies) consisting of pairs of attributes (e.g., a reduction of $x\%$ in the annual risk of death for $\pounds y$). Ratings of this sort are useful if the tradeoffs are independent of other available options (e.g., “rationally irrelevant” factors like the range of values on each attribute). If a change from 30 to 60 min is worth a change from $\pounds 10$ to $\pounds 20$, then this should be true regardless of whether the range of available monetary options is from $\pounds 10$

to $\pounds 20$ or from $\pounds 1$ to $\pounds 200$. Utility of each option should depend on its consequences, not on what other options are considered; yet, such independence is often not found (Baron, 1997).

ANOMALIES IN PSYCHOPHYSICS, CHOICE, AND VALUATION

Preferences between pairs of options may often be reliable when the options are of the same qualitative type and/or differ on a single dimension. But interesting choices tend to be more challenging in two ways: they typically involve trading off between different dimensions and comparing qualitatively different types of outcome, which is difficult even on a single dimension (such as when “comparing apples and oranges”). This article focuses on explaining one prominent psychological anomaly, *trade-off inconsistency* (TI), which violates the independence assumption of neoclassical economics. To illustrate TI, note that a person may easily judge that eating marginally more calories is preferable to eating slightly fewer; and marginally less risk of diabetes is preferable to more. But deciding what to eat involves trading off between these, and other dimensions against factors such as the pleasure obtained from food; and people appear systematically inconsistent in making such tradeoffs.

The basic underpinnings of this cognitive limitation can be found in psychophysics. Psychophysical results indicate that people do not have access to mental representations of “absolute” magnitudes, at least for perceptual stimuli; and hence base their decisions on relative, not absolute, values. For example, Garner (1954) asked people to choose a tone half-as-loud as a comparison tone. However, one group of people received candidate tones that included the half-as-loud tone but were mostly quiet, while another group received tones that also included half-as-loud tone but were mostly loud. In both groups, the recipients selected a tone in the middle of the range, so that the “quiet” group’s estimates of the half-loudness were much lower

than the “loud” group. The conclusion was that people have no grasp on absolute loudness; instead, they are more influenced by the alternative choice options and scarcely at all by the comparison stimulus.

Inspired by Garner’s (1954) study and similar studies in psychophysics where it was shown that people have no grasp on absolute loudness (Laming, 1997), Stewart et al. (2003) showed that such psychophysical principles carry over to risky choice, where the option set (i.e., the context) affects peoples’ choices, because there is no fixed internal scale according to which people make their judgments of the values of certain options. In particular, when participants choose to trade-off risk and monetary return by choosing a gamble (“ p chance of x ”) from a varying range of options/gambles, the range (full range of options, only safe options, and only risky options) was found to almost completely determine the choice: people chose based not on absolute risk–return level, but on the risk–return level relative to the other available gambles. In parallel work on risky financial decisions, Vlaev et al. (2007a,b) found similar effects of skew and range, in line with the range–frequency theory of magnitude judgment (Parducci, 1965, 1995). In particular, the range of options offered as potential saving levels and investment risks largely influenced the selected options, and the rank of riskiness of the investments affected the preferences for risky investment such that options with higher rank were considered as more risky and unattractive. Vlaev and Chater (2006, 2007) report similar relativistic effects in strategic decision making, where people do not have absolute grip of the level of cooperativeness implicit in each social dilemma game and, instead, such games are assessed, and strategic choices are made, relative to the range, rank, or mean cooperativeness of previous games that have been encountered. Such relativistic responses are inconsistent with an absolute measure of utility, or related concepts such as the value-function in prospect theory (Kahneman and Tversky,

1979) and rank-dependent utility (Quiggin, 1993), because these TIs cause preference reversals, which cannot be explained by classical, reference-, or rank-dependent utility models.

These results are related to other studies of context effects. Baron (1997) presents evidence that people judge the utility of a change or a difference as a proportion of the overall magnitude, even when the change alone is more closely related to their goal; and as a result, judgments are dependent on the maximum magnitude on each attribute scale. Another classic example is the jacket-calculator problem of Tversky and Kahneman (1981, p. 457; replicated by Darke and Freedman, 1993), which asks people to make a hypothetical choice between a jacket for \$125 and a calculator for \$15. When respondents were informed that the calculator is on sale for \$10 at another store located 20 min away, most of them preferred the trip to save the \$5, but only few respondents in another condition selected to make the trip to save \$5 on the jacket, although the real trade-off is about whether one would be willing to drive 20 min for \$5. Therefore, people judge the utility of saving \$5 as a proportion of the specific money attribute, not on the basis of a trade-off between money and time (opportunity cost); which goes against normative choice models such as multi-attribute utility theory (Keeney and Raiffa, 1976). Similar context effects are caused by relative reference points in studies of perception of economic attributes such as price (Niedrich et al., 2001).

Vlaev et al. (2009) demonstrate TI in an incentivized auction experiment (Becker-DeGroot-Marshack, 1964), in which participants choose to pay to avoid painful electrical shocks. These valuations were determined by two factors: recent intensities of other pains – medium pain provoked markedly higher price offers when it occurred with low pain rather than high pain; and immediately disposable income – higher offers were made when the endowment-per-trial was higher (i.e. individual ‘overall wealth’, which should not be affected by the small variations in the endowments per trial, was not an important factor) (see Figure 1). The estimated consumer-demand curves for pain relief, which indicate the quantity of pain relief expected to be bought at different prices, also exhibited relativistic patterns – higher demand for relief from medium pain when paired with low rather than high pain. This

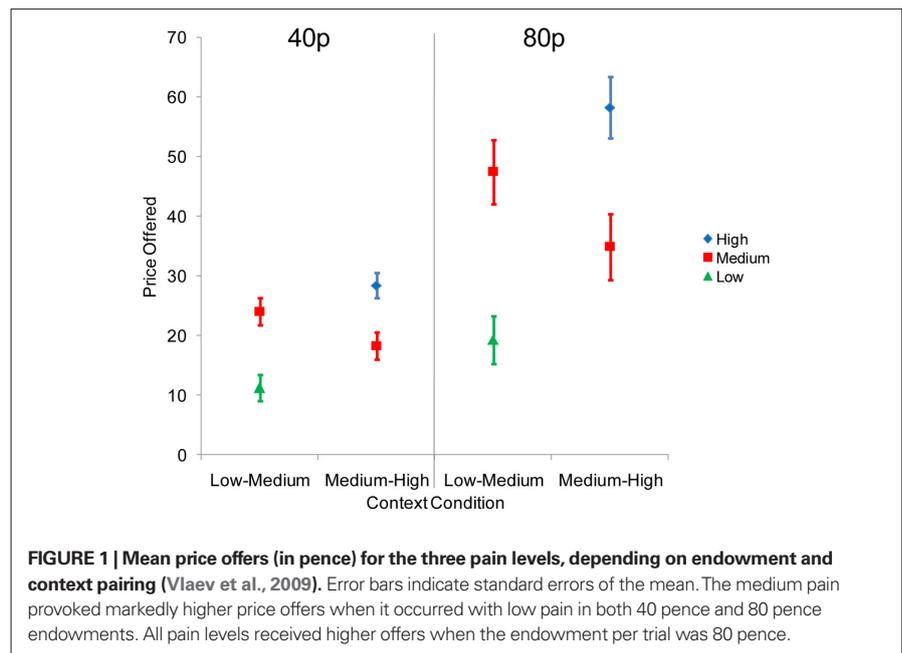


FIGURE 1 | Mean price offers (in pence) for the three pain levels, depending on endowment and context pairing (Vlaev et al., 2009). Error bars indicate standard errors of the mean. The medium pain provoked markedly higher price offers when it occurred with low pain in both 40 pence and 80 pence endowments. All pain levels received higher offers when the endowment per trial was 80 pence.

suggests that the subjective value attributed to pain relief is remarkably malleable and people cannot establish a stable trade-off between money and pain (note that stable trade-off is assumed in economic valuation of pain, which informs the market price of analgesics, the cost-effectiveness of clinical treatments, compensation for injury, and the response to public hazards). Ariely et al. (2003) demonstrated similar reference-dependence of preferences by showing that willingness-to-pay to avoid aversive stimuli is strongly biased toward arbitrary price anchors.

Another type of TI is due to variation in decision ‘content’ (different from ‘context’ or the choice set), which produces variation in preferences for risk, because people’s knowledge of event frequencies ‘leaks’ into decisions even when event likelihood information is explicitly provided (Kusev et al., 2009).

In summary, the fact that individuals are subject to contextual biases, and as a result behave inconsistently, is important. From this basis, this article aims to offer an explanatory account of such TI effects.

COMMENSURABILITY EXPLAINS INCONSISTENT DECISION MAKING

Trade-off inconsistency are not cognitive oddities – they arise systematically from basic properties of the cognitive system. A fundamental cognitive principle, which promises to explain the above phenomena,

is *qualitative incommensurability* (QI). This principle postulates that people are unable to systematically compare qualitatively different options or outcomes on a single value dimension – an assertion based on existing evidence that such comparisons are extremely difficult (Luce and Green, 1978; Stewart and Chater, 2003). Recent psychophysical research has suggested that basic judgments of perceptual quantities like loudness, qualitatively different stimuli cannot be consistently compared, *even when judged on the same attribute*, and judgments are distorted by the influence of previous context. Stewart and Chater (2003) presented participants with two different auditory stimuli, either pure tones or white noise (‘buzzes’) delivered independently via headphones to each ear, and the task was to choose the loudest tone. The loudness of previous items strongly influenced the perceived loudness of the present item, but this effect was reduced when items were qualitatively different – a previous loud ‘tone’ made a present ‘tone’ quieter, but not a present ‘buzz,’ thus modifying the choice between them. Where this phenomenon occurs in relation to choices that are presumed to reveal preference (e.g., choice between using money to preserve a 80 acres of marshland vs. providing clean drinking water for 2000 people), we face a substantial puzzle, because preference reversals will be generated by varying the previous items (e.g., saving 180 acres of marshland vs.

clean drinking water for 1000 people). But then which choice reveals a person's "true" preference? There is no "neutral" context in which psychophysical, or other, judgments are made. Hence, this result raises concerns over the extent to which contextual effects undermine our ability to stably trade-off between qualitatively different goods.

Stewart and Chater's (2003) results in "perceptual choice" (see also Luce and Green, 1978) imply that even apparently similar dimensions (loudness of sounds) may not be commensurable. Psychophysical research also reveals that the decision context determining choice is often the immediately preceding stimuli, as in the related domain of absolute magnitude identification (Stewart et al., 2005).

In "preferential choice," most interesting decisions also involve comparing incommensurable properties – e.g., dietary restraint against risk of heart disease, mobility maintenance in patients against pain avoidance, financial vs. environmental factors, investment risk vs. return. Consider a person on a diet who may find it difficult to assess the pleasure they may get from either of two 100 calorie treats: a thin sliver of cake, or 10 grapes; likewise, a typical public policy budgeting decision involve relating qualitatively different outcomes, such as heart transplants, educational programs, or air quality improvements, against a fixed overall budget. These choices are difficult in part due to a lack of information (e.g., about how health behaviors relate to levels of health risk; or in public policy, the amount of marginal health, educational, or environmental benefit per pound spent). QI implies that the problem is more fundamental – even with perfect information, basic properties of the cognitive system appear to show that such preferences are undefined.

In contrast, normative theories and their descriptive deviations assume that an attribute's value is "translated" into a single underlying measure called "utility," which can be positive or negative; and after all attributes are independently "mapped" on this "common currency" scale, the overall utility of an option is determined by some additive process. To illustrate this point, imagine that a person is indifferent between eating an apple and two oranges, and she tries another extremely delicious fruit. Normative decision theories predict that this new experi-

ence should not affect her initial trade-off (e.g., by making her demand two apples for two oranges). Reference-dependent utility theories predict that the pleasure from these apples and oranges should be reduced, while the trade-off is preserved, which happens irrespective of whether she tries a delicious apple, orange, or any other fruit; because, the greater new utility affects all smaller utilities by changing their reference points (i.e., the pleasure from eating apples is not calculated on a separate utility scale). QI predicts that the trade-off is likely to vary depending on the commensurability between two options, and thus can differentially undermine the quality of choice. Therefore, the cases in which QI is strong enough to lead to context effects and preference reversals are the most intriguing.

To summarize, QI is a prominent psychological anomaly that has implications for explaining TI in choice and valuation across decision domains like health, the environment, finance, and consumer-spending. There are also crucial implications for normative theories of rational choice, consumer-theory (e.g., calculating a "commensurability index" between products before deriving their indifference curves), and the practical methodology of valuing non-market goods.

ACKNOWLEDGMENTS

The author would like to thank Nick Chater and Neil Stewart for their contribution in developing the ideas discussed in this article.

REFERENCES

- Anderson, N. H., and Zaluski, J. (1988). Functional measurement approach to self-estimation in multiattribute evaluation. *J. Behav. Decis. Mak.* 1, 191–221.
- Ariely, D., Loewenstein, G., and Prelec, D. (2003). Coherent arbitrariness: stable demand curves without stable preferences. *Q. J. Econ.* 118, 73–105.
- Baron, J. (1997). Biases in the quantitative measurement of values for public decisions. *Psychol. Bull.* 122, 72–88.
- Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behav. Sci.* 9, 226–232.
- Darke, P. R., and Freedman, J. L. (1993). Deciding whether to seek a bargain: effects of both amount and percentage off. *J. Appl. Psychol.* 78, 960–965.
- Garner, W. R. (1954). Context effects and the validity of loudness scales. *J. Exp. Psychol.* 48, 218–224.
- Green, P. E., and Srinivasan, V. (1990). Conjoint analysis in marketing: new developments with implications for research and practice. *J. Mark.* 45, 33–41.
- Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291.

- Keeney, R. L., and Raiffa, H. (1976). *Decisions With Multiple Objectives: Performances and Value Trade-Offs*. New York: Wiley.
- Kusev, P., van Schaik, P., Ayton, P., Dent, J., and Chater, N. (2009). Exaggerated risk: prospect theory and probability weighting in risky choice. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 1487–1505.
- Laming, D. (1997). *The Measurement of Sensation*. Oxford: Oxford University Press.
- Louviere, J. J. (1988). *Analyzing Individual Decision Making: Metric Conjoint Analysis*. Newbury Park, CA: Sage.
- Luce, R. D., and Green, D. M. (1978). Two tests of a neural attention hypothesis for auditory psychophysics. *Percept. Psychophys.* 23, 363–371.
- Niedrich, R. W., Sharma, S., and Wedell, D. H. (2001). Reference price and price perceptions: a comparison of alternative models. *J. Consum. Res.* 28, 339–354.
- Parducci, A. (1965). Category judgment: a range-frequency theory. *Psychol. Rev.* 72, 407–418.
- Parducci, A. (1995). *Happiness, Pleasure, and Judgment: The Contextual Theory and its Applications*. Mahwah, NJ: Erlbaum.
- Quiggin, J. (1993). *Generalized Expected Utility Theory: The Rank-Dependent Model*. Boston: Kluwer Academic.
- Stewart, N., Brown, G. D. A., and Chater, N. (2005). Absolute identification by relative judgment. *Psychol. Rev.* 112, 881–911.
- Stewart, N., and Chater, N. (2003). "No unified scales for perceptual magnitudes: evidence from loudness," in *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*, eds R. Alterman and D. Kirsh (Boston, MA: Cognitive Science Society), 1116–1121.
- Stewart, N., Chater, N., Stott, H. P., and Reimers, S. (2003). Prospect relativity: how choice options influence decision under risk. *J. Exp. Psychol. Gen.* 132, 23–46.
- Tversky, A., and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science* 211, 453–458.
- Vlaev, I., and Chater, N. (2006). Game relativity: how context influences strategic decision making. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 131–149.
- Vlaev, I., and Chater, N. (2007). Context effects in games: local versus global sequential effects on choice in the prisoner's dilemma game. *Judgm. Decis. Mak.* 2, 380–389.
- Vlaev, I., Chater, N., and Stewart, N. (2007a). Financial prospect relativity: context effects in financial decision making under risk. *J. Behav. Decis. Mak.* 20, 273–304.
- Vlaev, I., Chater, N., and Stewart, N. (2007b). Relativistic financial decisions: context effects on retirement saving and investment risk preferences. *Judgm. Decis. Mak.* 2, 292–311.
- Vlaev, I., Seymour, B., Dolan, R., and Chater, N. (2009). The price of pain and the value of suffering. *Psychol. Sci.* 20, 309–317.

Received: 18 July 2011; accepted: 11 October 2011; published online: 15 November 2011.

Citation: Vlaev I (2011) Inconsistency in risk preferences: a psychophysical anomaly. *Front. Psychology* 2:304. doi: 10.3389/fpsyg.2011.00304

This article was submitted to *Frontiers in Cognition*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Vlaev. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.