Supplement to "Contextual Effects in Salary Satisfaction"

Michael H. Birnbaum and Julien Rouvere[1]

[1]California State University, Fullerton

July 8, 2023

Abstract

This supplement contains (1) A figure of the Beta distributions estimated to represent the average effective contexts of 4 groups of participants who worked full time and who reported different levels of income. (2) An analysis that assumes that the prior context and the experimental context combine by vertical averaging. That is, the cumulative distribution that reflects the effective context is modeled as an average of the cumulative distribution of experimental stimuli and the cumulative residual context, which is estimated as a Beta distribution. (3) A figure of the fit of this theory to mean judgments by groups who reported different levels of income and who received either the C1 or C2 cubic distributions in Experiment 1. (4) This supplement contains a table recommended by a reviewer that may be helpful for reminding the reader of the key ideas of the different theories and for summarizing the six theories of contextual effects to students.

Residual Context Effects: The Estimated Effective Contexts

This section is partially redundant with the main paper, but it adds a figure of the estimated effective contexts for different groups of people who had different incomes. Equations from the main text are also included here, to facilitate comparison with the following section, which presents a model in which the effective context is the vertical average of the residual (prior) context and the experimental context.

To examine the relationships between income and judgments, we divided data for the 191 participants in Experiment 1 who reported working full time (38-42 hours per week) into four groups according to self-reported income. This analysis included both main and unusual data and combines across experimental contexts. There were 48, 36, 48, and 59 individuals who had incomes less than \$40K, \$40K to \$52K, between \$52K and \$85K, and \$85K and above, respectively.

The mean judgments in Figure 4 of the main paper were fitted with the assumptions that $s = u(x) = x$, $w = 0.5$, and that the average effective context is distributed as a Beta distribution with parameters that depend on a group's income level. The data were fit to the equation:

$$P_g(x) = 6[wB(x, \alpha_g, \beta_g, y_{0g}, y_{mg}) + (1 - w)\frac{(x - y_{0g})}{(y_{mg} - y_{0g})}] + 1 \quad (1)$$

where $P_g(x)$ is the predicted mean judgment of salary x by income Group g; $B()$ is the cumulative Beta distribution; α_g and β_g are the estimated shape parameters for the Beta distribution in Group g; y_{0g} and y_{mg} are the estimated minimum and maximum in the effective context for Group g; that is, these are stimuli that would have been judged 1 and 7, respectively.

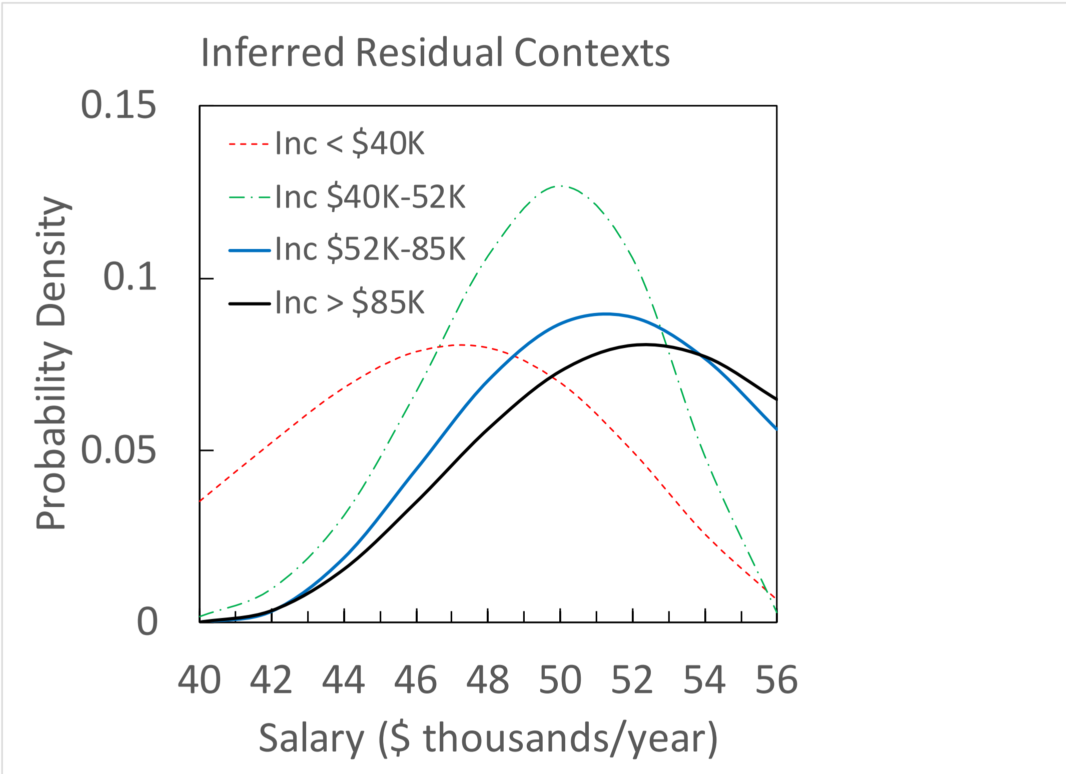

Figure 1: Estimated Beta distributions representing the effective contexts for groups with different incomes, averaged over C1 and C2 experimental contexts. The red, dashed curve is for people with incomes less than $40K; the green, dash-dotted curve is for incomes from $40K to $52K; the solid, blue and black curves are for those with incomes above $52K to $85K, and above $85K, respectively.

For groups with lowest to highest incomes, respectively, least-squares estimated minima were $26.94, $35.31, $39.89, and $39.35 thousand; estimated maxima were $58.28, $56.77, $67.32, and $69.15 thousand, respectively. The estimated shape parameters for the Beta distribution were $(\alpha, \beta) = (5.99, 3.72), (6.91, 3.67), (4.20, 5.53)$, and $(4.25, 5.18)$, respectively. These are single-peaked, bell-shaped distributions that, as one might expect, shift to the right as income increases.

Figure S1 shows the estimated residual contexts for these four income groups in Experiment 1. The peaks shift to the right as income level increases. In addition, the variance of the estimated distribution for those with incomes between $40K and $52 (the range of experimental salaries in Experiment 1) has smaller variance, so the cumulative distribution (and consequently the predicted judgments) are steeper over this range for these participants.

In this curve fitting, the estimated "effective" minima and maxima are now estimated parameters (instead of the actual minima and maxima controlled by the experimenter), so they can fall outside the actual range of the stimuli used in the experiment. Their estimation depends crucially on the assumed Beta distribution used to extrapolate to their values. Therefore, although this fitting method gives a good reproduction to these data and we think that these estimated parameters might be used to predict new results on the same range for the same income groups, we advise caution in extrapolating predictions outside the range of salaries actually used in the study.

Fitting RF Model Combining Experimental and Residual Contexts

In order to fit RF theory to the effects of both the income level and the experimentally manipulated frequency distributions, we subdivided the data for those who worked full time into those who received Conditions C1 or C2 and those who earned less than or equal to $52K ("low" income) and those who earned more than $52K ("high" income). That resulted in four groups with 66, 67, 31, and 49 participants who reported low salaries in C1 and C2 or high salaries in C1 or C2, respectively. The mean judgments for these four groups of participants are shown in Figure S2 as unfilled circles, filled squares, unfilled triangles or filled diamonds, respectively. The data show the double crossover in each income group, the same as was shown in the main paper, averaged over income.

Figure 2: Fit of Range-Frequency theory to experimentally manipulated frequency distributions C1 and C2 for groups with incomes below and above $52K. The model assumes that effective context is the average of residual context and experimental context.

An extended RF model was fit to the data as follows:

$$P_{kg}(x) = 6[w(aF_k(x) + (1-a)B(x, \alpha_g, \beta_g, y_{0g}, y_{mg})) + (1-w)\frac{(s - s_{0g})}{(s_{mg} - s_{0g})}] + 1 \qquad (2)$$

where $P_{kg}(x)$ is the predicted judgment of Salary x in experimental Context k by income Group g; w is the weight of the effective frequency distribution, which itself is assumed to be a vertical average of the cumulative experimentally manipulated frequency in Context k, F_k, and the residual context, represented by $B()$, a cumulative distribution with shape parameters that depend on income; and a is the weight of the experimental cumulative frequency relative to the cumulative residual context.

For income Group g, the residual context is approximated by a cumulative Beta distribution with four free parameters, $\alpha_g, \beta_g, y_{0g}$, and y_{mg}. As in Equation 9 of the main paper, α_g and β_g are the Beta function shape parameters; y_{0g} and y_{mg} are the minima and maxima of the residual distribution for Group g. It was assumed for simplicity[1] that $s = u(x) = x$, for this small range of x, so $s_{0g} = y_{0g}$ and $s_{mg} = y_{mg}$. There are therefore 8 parameters for the residual contexts in two income groups plus 2 (for w and a) to estimate from the 28 mean judgments.

The least-squares estimates of α, β, y_{0g}, and y_{mg} were 96.2, 76.7, \$0, and \$87.0 thousand for the low income group; and they were 10.3, 44.9, \$40 thousand, and \$91.5 thousand, for the high income group, respectively. The estimated weights were $a = 0.50$ and $w = 0.69$. The sum of squared deviations was 0.333.

As shown in Figure S2, the model does a fairly good job in reproducing the actual mean judgments. Therefore, we can retain the RF theory with the assumptions that the residual contexts can be fit by Beta distributions whose parameters depend on the income of the participants, and that the effective context is a vertical average of the experimental and prior contexts.

Although this expanded RF theory fits the data of Experiment 1 pretty well, it should also be clear that we have used a relatively large number of parameters with flexible Beta distributions to fit the data.[2] It would be useful to see better tests of the vertical averaging assumption for the combination of residual and experimental contexts, with a greater number and variety of experimental contexts.

[1] We considered the experimental design of Experiment 1, with a small fixed range of salaries and only two conditions, to provide insufficient constraint to distinguish the shape of the $u(x)$ function from the shape of the prior contexts.

[2] Experiment 2 used mostly lower division college students who had very little variation in income, so Experiment 2 could not be used to analyze residual contexts.

A reviewer recommended that we expand the section on model fitting and include a comparison of model fits in the main paper. However, we decided to relegate this section on model fitting of residual and experimental contexts to the Supplement because we do not want to distract the reader from our use of critical tests (rather than model-fitting and indices of fit) to decide among the models of contextual effects. The main reasons to prefer RF theory to the alternatives are (1) that RF theory can account for the double crossover when the experimental contexts have specific cubic distributions and (2) RF theory correctly accounts for the effects of manipulating the endpoints of the distribution.

1 A Table Summary of 6 Theories

A reviewer recommended that this table could be useful as an adjunct to this paper, to help a reader keep in mind the main distinctions among the theories of contextual effects. It should also be useful in teaching these theories to students.

References

Helson, H. (1964). *Adaptation-Level theory.* Oxford, England: Harper & Row. https://psycnet.apa.org/record/1964-35039-000

Johnson, D. M., & Mullally, C. R. (1969). Correlation-and-regression model for category judgments. *Psychological Review, 76*(2), 205–215. https://doi.org/10.1037/h0027227

Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review, 72*(6), 407–418. https://doi.org/10.1037/h0022602

Putnam-Farr, E., & Morewedge, C. K. (2021). Which social comparisons influence happiness with unequal pay? *Journal of Experimental Psychology: General, 150*(3), 570-582. https://doi.org/10.1037/xge0000965

Stewart, N., Chater, N., & Brown, G.D.A. (2006). Decision by sampling. *Cognitive Psychology, 53*, 1-26. https://doi.org/10.1016/j.cogpsych.2005.10.003

Wort, F., Walasek, L., & Brown, G. D. A. (2022). Rank-based alternatives to mean-based ensemble models of satisfaction with earnings: Comment on Putnam-Farr and Morewedge (2020). *Journal of Experimental Psychology: General, 151* (11), 2963–2967. https://doi.org/10.1037/xge0001237

Table 1: Theories of Contextual Effects

Abbrev	Brief summary
AL	Helson (1964): In AL theory, the response to a stimulus depends only on its relation to the AL., which is the stimulus whose subjective value is a weighted average of the subjective values of all stimuli presented or remembered from previous experience.
CR	Johnson and Mullally (1969) proposed that the response to a stimulus is a linear function of the distance of the stimulus from the mean, relative to the spread of the stimuli.
ID	Wort, Walasek, & Brown (2022) theorized that people infer a normal distribution from the mean and endpoints of salaries presented and base judgments on ranks implied by the inferred distribution.
EN	Ensemble Theory: Putnam-Farr and Morewedge (2021) wrote, "A person making an above average salary would then compare her salary to the group mean and highest salary, for instance, whereas a person making a below average salary would compare his salary to the group mean and lowest salary.... our ensemble representation account implies that people should be insensitive to other properties of groups, ... such as their relative rank in the group."
DbS	Stewart, Chater, and Brown (2006) proposed that the response to a stimulus is a function of the rank of a stimulus in the distribution of stimuli sampled from memory. When comparing two stimuli, people only make ordinal comparisons.
RF	Parducci's (1965) Range-Frequency theory assumes judgment of a stimulus is a compromise between the position of a stimulus relative to the endpoints and the cumulative frequency (rank) of the stimulus in the distribution of stimuli.