The Effects of Sample Size on Different Measures of Subjective Correlation

Justin M. Gilkey, Richard B. Anderson, and Michael E. Doherty

Abstract

The ability to accurately infer population correlations is important in our everyday lives. One factor influencing the evaluation of population correlations is the size of the sample used in making the inference. Typically, subjective correlation has been measured using a rating, prediction, or frequency estimation task. Recent studies evaluating the impact of sample size on subjective correlation suggests that the effect of sample size on subjective correlation may be impacted by the type of response task used. Controlling for stimulus and procedural differences, the results of this study suggest that there might be important differences between tasks which measure subjective correlation directly, and those which measure subjective correlation indirectly.

Introduction

When judging the relationship between events in the environment, people typically must rely on sample information drawn from the environment to make their judgment. Of obvious importance to drawing inferences from sample data is the number of observations contained in the sample.

Over the past decade the relationship between sample-size and subjective correlation has become an increasingly active area of research. The existing literature is inconsistent regarding the impact of sample size on the perceived correlation between two variables. When assessed by a rating task (Clement, Mercier, & Pasto, 2002) or by an estimation task (Anderson, Doherty, & Gilkey, 2006), subjective correlation has increased with sample size. When measured by a prediction task (Kareev, Lieberman, & Lev,

1997) subjective correlation has decreased as sample size increases.

The present study is designed to investigate performance on these tasks within a single study, thereby controlling for extraneous factors as alternative explanations for performance differences across the three tasks. Figure 1 details the predictions.

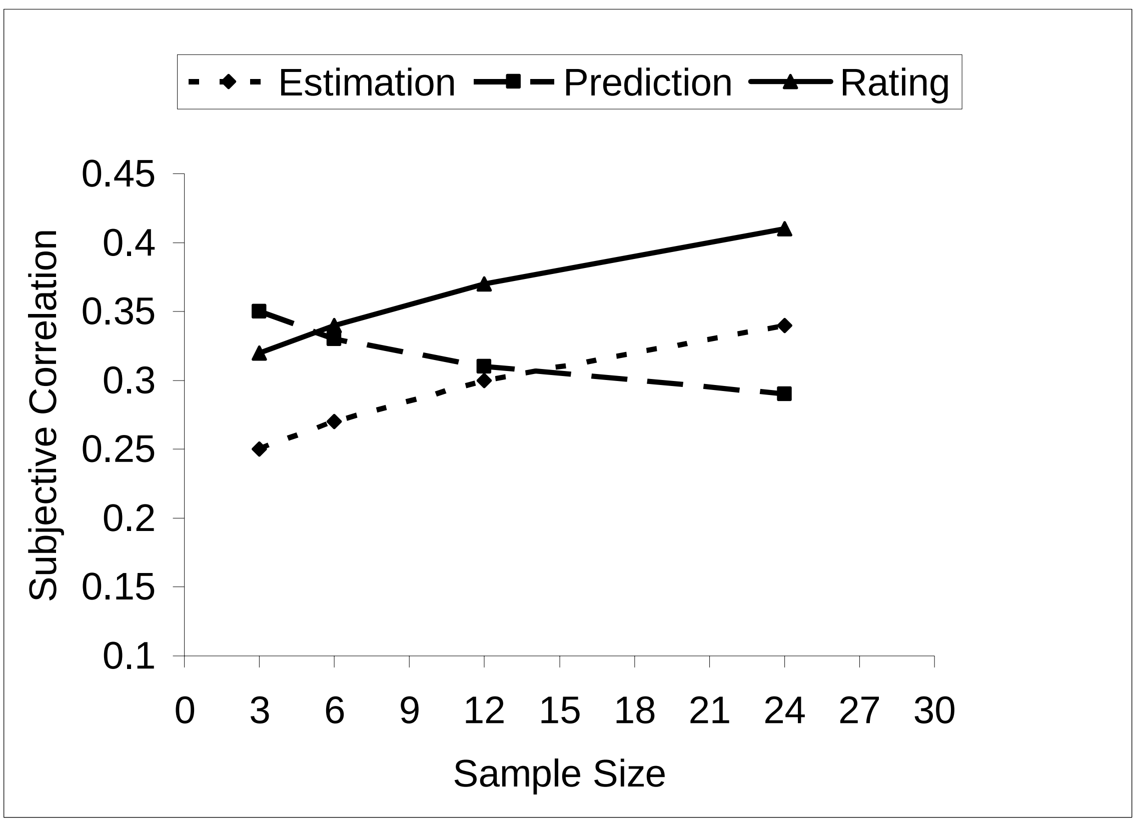

Figure 1. Predicted results of subjective correlation as a function of sample size aggregated across 3 levels of population correlation

Method

Fifty-Seven participants (23 male, 34 female) were recruited from undergraduate psychology courses for the current study. Each participant was tested individually on a computer running Windows XP. All experimental applications were written with e-Prime experiment development software. The stimuli for the experiment consisted of drawings representing envelopes and payment statements. The envelopes were either brown or white, and the payment statements indicated either credit card or check payments.

Participants were randomly assigned to one of three dependent measure groups (prediction, estimation, or rating). On each of 93 trials, participants were serially presented with 3, 6, 12, or 24 envelopes; after a 500 MS delay the contents of the envelope were displayed to the right of the envelope (Figure 2 and 3). The envelopecontent pairs were randomly drawn from populations with correlations of 0, .4, or .8.

Thus the study utilized a 3×4×3 mixed factorial design in which participants were randomly assigned to one of three dependent variable conditions (rating, estimation, and prediction). Both sample size (3, 6, 12, and 24) and population correlation (0, .4, and .8) were manipulated within subjects.

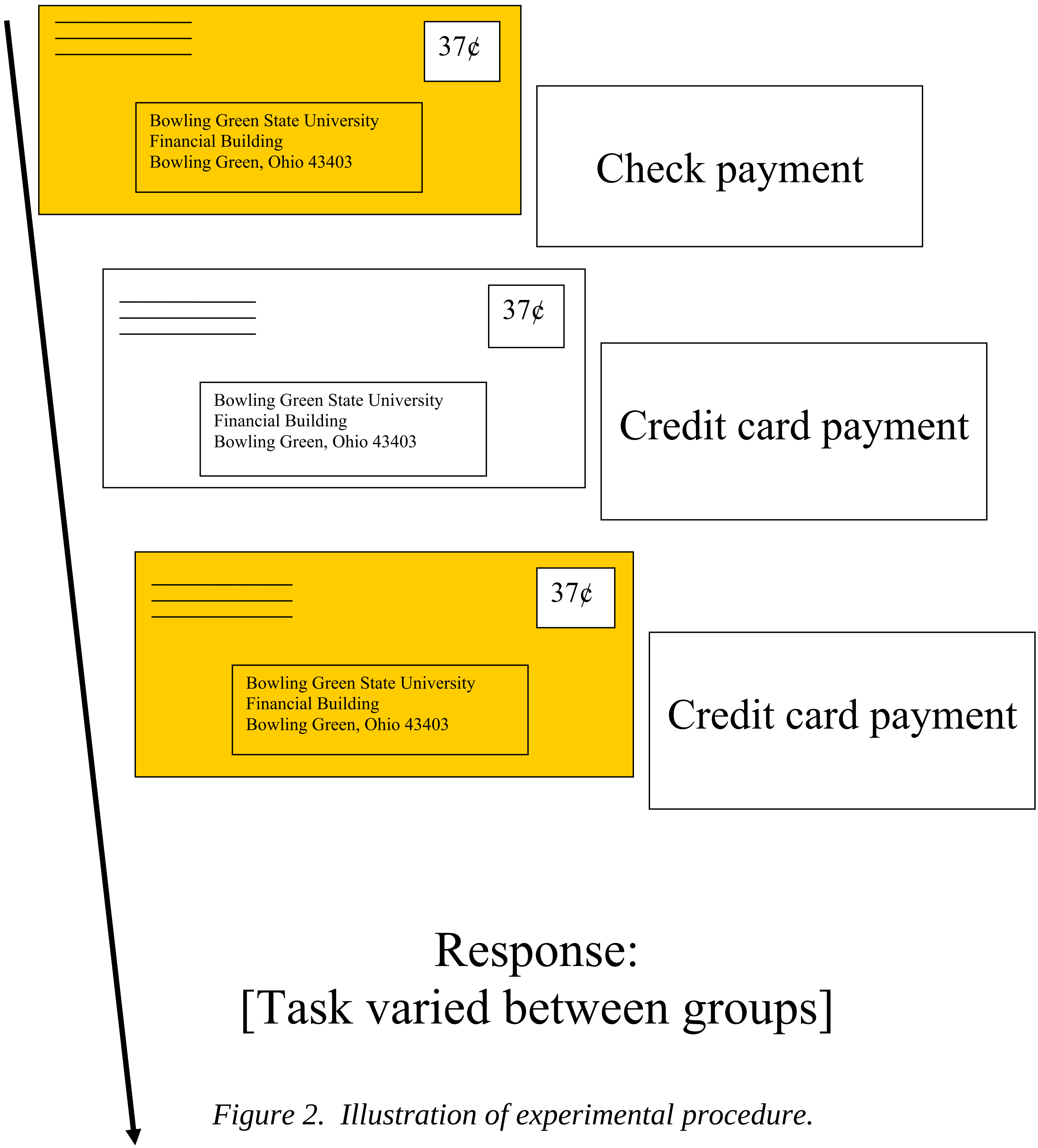

Figure 2. Illustration of experimental procedure.

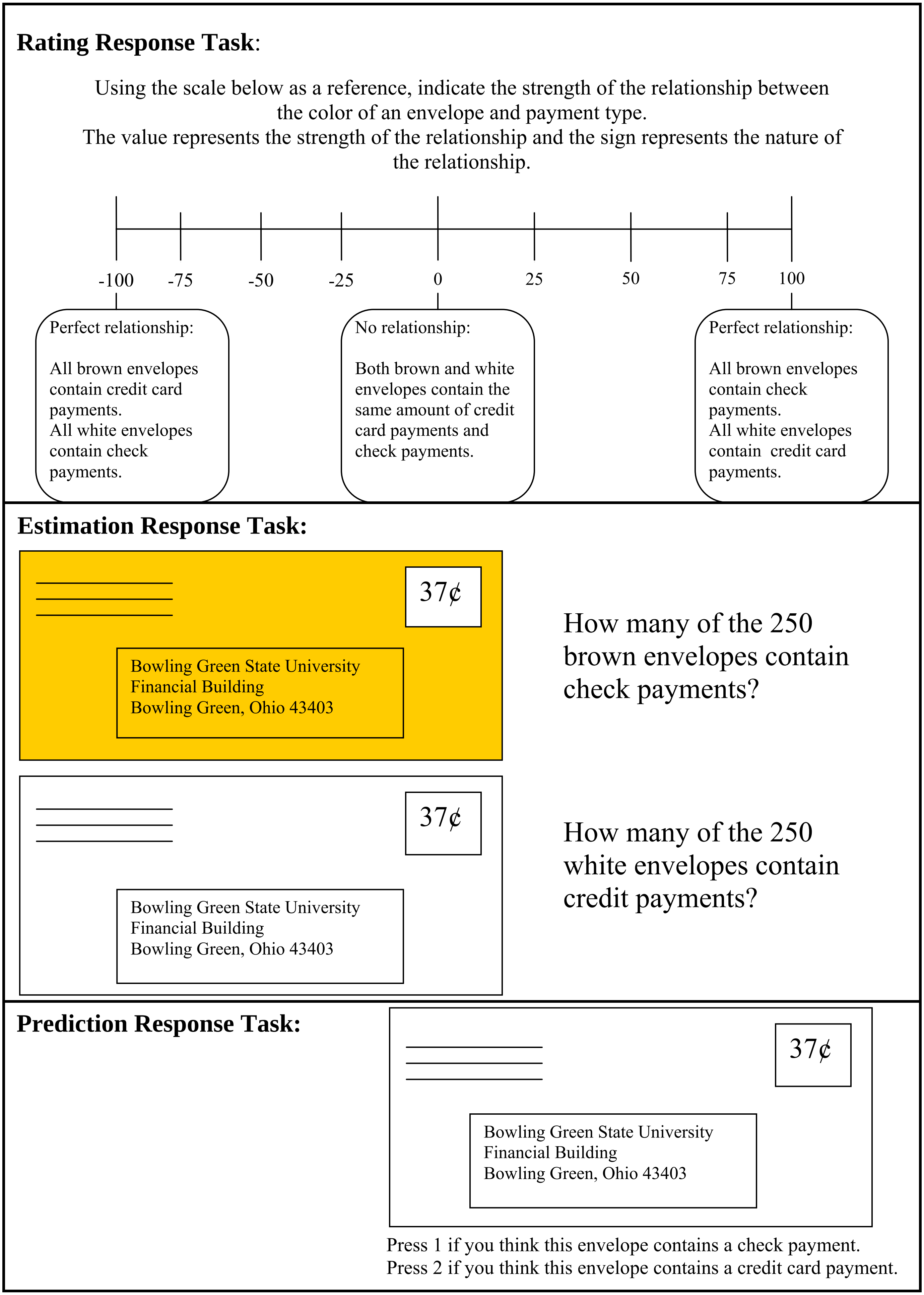

Figure 3. Illustrations of response tasks.

Results

Subjective correlations were computed using methods appropriate for each response task. For the rating task, a participant's rating of the strength of the relationship was directly used as a measure of subjective correlation. While both the estimation and rating task relied on the phi coefficient, φ, which is identical to Pearson's *r* for two discreet variables:

$$\varphi = BC-AD/\sqrt{[(A+B)(C+D)(A+C)(B+D)]}$$

where A, B, C, and D refer to the frequencies of occurrences of the four possible variable-level combinations of envelope color (brown or white) and payment type (credit card and check) .

Due to an insufficient number of participants, the results for each of the response tasks have been aggregated across the three levels of population correlation. See Figure 4.

Overall there was a significant main effect of response task. $F(2,54) = 4.096$, p=.022.

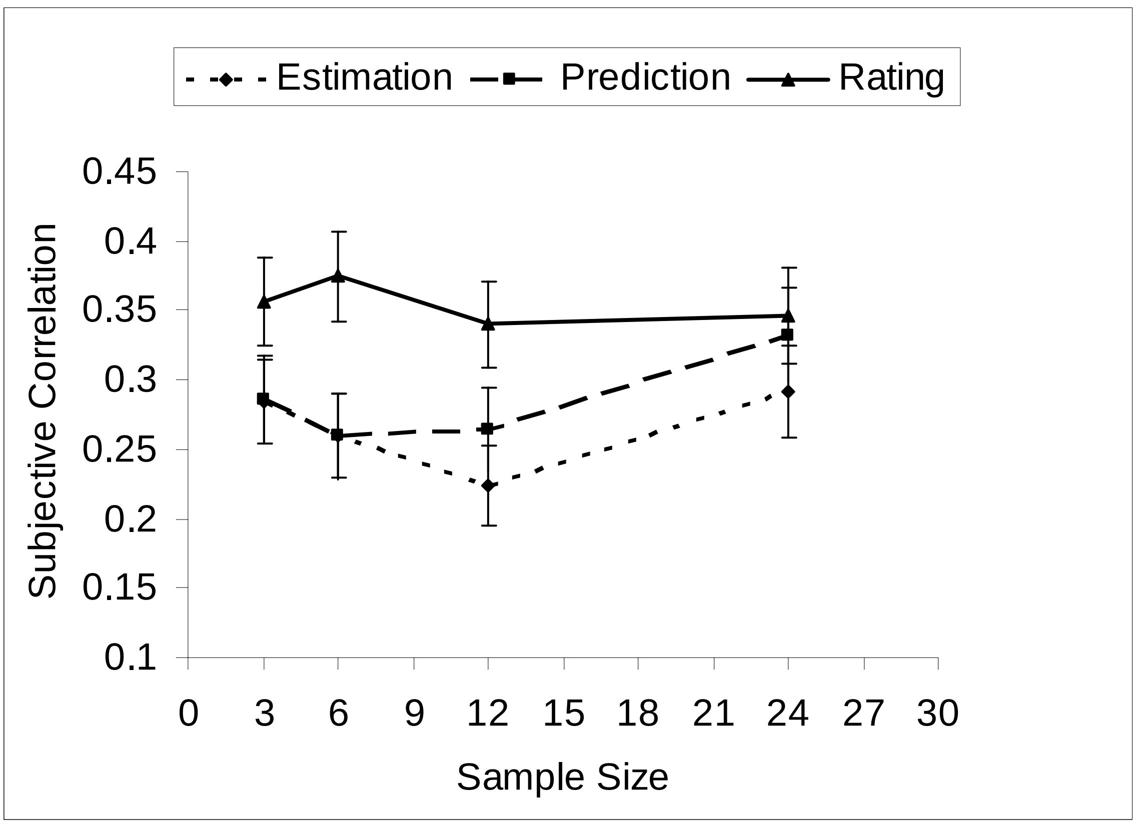

Figure 4. Subjective correlations as a function of sample size for three response types.

Discussion

The results of the study show the general pattern that the means of the rating task tended to be closer to the objective correlation, .4 in this case, than either the estimation or prediction task. This may be related to the fact that the rating task is arguably a more direct assessment of perceived correlation than are the other tasks. In addition, when computing the subjective correlations for participant's responses in the indirect tasks (cell frequency estimation and stimulus prediction), it is possible for the derived estimate to be undefined. In contrast the rating task, does not allow for undefined subjective correlations.

Sampling theory predicts that randomly drawn samples are more likely to have an undefined correlation as the size of the sample decreases (Anderson, Doherty &

Friedrich, 2005). Table 1 confirms this prediction. Table 1 also indicates that subjective correlations follow the same pattern as sample correlations and that the presence or absence of an undefined subjective correlation may be correlated with the presence or absence of an undefined sample. Thus trials which subjective correlations are undefined are related to trials where participants' experience samples which have an undefined correlation.

Overall the results suggest that participants' perceived correlations between two variables, are closer to the objective correlation when measured directly as opposed to indirectly. This difference could be related to the possibility of deriving undefined subjective correlations from participants' responses.

Table 1. Frequencies of undefined sample correlations ($U_{samp.}$), undefined subjective correlations ($U_{subj.}$), and the correlation between defined/undefined sample correlation and defined/undefined subjective correlation (R_u) for each response task at each level of sample size ($N_{samp.}$)

Task	$N_{samp.}$	$U_{samp.}$	$U_{subj.}$	R_u
Estimation				.246
	3	138	26	
	6	18	3	
	12	1	0	
	24	0	0	
Rating				----
	3	147	----	
	6	22	----	
	12	0	----	
	24	0	----	
Prediction				.187
	3	140	13	
	6	17	4	
	12	0	2	
	24	0	2	

References

Anderson, R. B., Doherty, M. E. & Gilkey, M. J. (2006). Effects of Sampling Ecology and Correlation Judgment. Presented at Cognitive Science Society.

Anderson, R. B., Doherty, M. E., & Friedrich, J. C. (2005). *Statistical / Ecological Factors in the Effects of Sample Size on Correlation Inference*. Presented at Psychonomic Society.

Clement, M., Mercier, P., & Pasto, L. (2002). Sample size, confidence, and contingency judgment. *Canadian Journal of Experimental Psychology, 56*, 128-137.

Kareev, Y., Lieberman, I., & Lev, M. (1997). Through a narrow window Sample size and the perception of correlation. *Journal of Experimental Psychology: General, 126*, 278-287.

Acknowledgments

This study was partially supported by a grant from the National Science Foundation.

Contact

Justin M. Gilkey <Jgilkey@bgsu.edu>

Richard B. Anderson <Randers@bgsu.edu>

Michael E. Doherty <Mdoher2@bgsu.edu>

Reprint requests