What Makes Tangible Incentives Motivating?
Development of Justifiability and Social Utility Scales

Scott A. Jeffrey
Leon Hess Business School, Monmouth University

Author Note

Correspondence concerning this article should be addressed to Scott A. Jeffrey, Monmouth University, Leon Hess Business School 400 Cedar Avenue West Long Branch, NJ 07764. Email: sjeffrey@monmouth.edu

Acknowledgement: The authors would like to thank Monmouth University for funding this research.

Ethics approval: This research received IRB approval from Monmouth University in Application SP1476

Abstract

In two surveys of working adults, the authors develop scales for the latent variables of justifiability and social utility. Statistical analysis shows that the statements developed provide good measurement of these variables. Using these scales will help researchers understand how tangible incentive motivate employees. In addition, firms can use employee responses to these statements to optimize the design and implementation of their incentive programs

Keywords*: Tangible Incentives, Scale Development, Motivation.*

Introduction

Tangible non-cash incentives such as travel, merchandise, and gift cards, are extensively used by organizations. In 2015, 84% of firms used some form of this type of incentive and the total spend was almost $90 Billion US, up from $61.4 Billion in 2013 ("Incentive Marketplace Estimate Research Study | Research | The Incentive Research Foundation," 2016). There is a significant body of research that shows that in many conditions, tangible rewards perform better that the equivalent value of cash as an incentive (Jeffrey, 2009; Jeffrey & Adomdza, 2011; Shaffer & Arkes, 2009). Thus, we know that tangible incentives can be motivating. However, we do not yet fully understand how to measure the mechanisms that make them motivating. This article hopes to begin addressing that gap by creating scales to measure motivational properties.

In a theoretical paper, Jeffrey and Shaffer (2007) outlined four mechanisms that they believed would cause tangible incentives to be motivating: Justifiability, Social Reinforcement, Evaluability, and Separability. While all four constructs contribute to the motivational power of tangible incentives, we only develop scales for the constructs of justifiability and social reinforcement in this article. We leave out the other two because past research has shown them to be adequately assessed with one or two direct questions and so do not need to be treated as a latent variable (Jeffrey, 2003; Jeffrey & Adomdza, 2011).

Justifiability

One feature of many tangible non-monetary incentives is that they are items that the recipients often see as luxuries; things that employees could not normally justify buying for themselves, even if they had sufficient funds. If an item is something that an employee values highly but would never purchase on his or her own, then the opportunity to earn it as a reward for hard work provides a way to obtain the desired object without violating one's standards of justification (Hsee, 1996). For example, a salesperson might never propose that his or her family take an expensive and "frivolous" trip to Hawaii, but everyone might be pleased if it were earned as a reward for hard work. If the employee is constrained to accept the award available, there is no need to justify its consumption, causing hard work to become an attractive way to acquire a good that would be difficult to obtain through other means. In this case, a tangible incentive provides a guilt-free (or at least a reduced guilt) way to obtain something valuable, thereby increasing its value.

Social Reinforcement

One of the most important rewards for a job well done is the acknowledgement of one's performance by peers, supervisors, family, and friends. This social utility comes from others knowing about the good performance rather than the receipt of an incentive per se. This is one of the central tenets behind most reinforcement theories of motivation (Alport, 1954; Bandura, 1969; Luthans & Stajkovic, 2000; Mahoney, 1974; Stajkovic & Luthans, 1997).

A tangible non-monetary incentive is an effective carrier of this social utility because this type of award is a visible sign of performance that others will know about, making it unnecessary for the employee to advertise that he or she earned the incentive. Many people feel that it is inappropriate to directly brag about their good performance (Scopelliti, Loewenstein, & Vosgerau, 2015), but would feel free to talk about golf clubs or a vacation they received from the firm for good performance. By providing a less proscribed means to bring attention to one's good performance, this social utility enhances the value of a tangible non-monetary incentive.

Method

We follow the recommended best practices as outlined by Wright, Quick, Hannah, and Hargrove (2017) and DeVellis (2012). First, the authors of this article sat together to discuss the constructs under study and define the important concepts in each latent variable. From there, we designed statements that we felt would address those underlying constructs. We wrote the survey using Qualtrics™ and directed recruits from Amazon's Mechanical Turk (MTurk) to a university server where the survey was hosted.

Study 1

For study one, we developed six statements for justifiability and five statements for social reinforcement. For justifiability we used the three statements from Jeffrey (2003) with additional statements created by the authors. For social utility, the authors used the three statements used in past research (Jeffrey and Adomdza, 2009; Jeffrey, 2003), and created two additional statements.

We provided a background story to participants prior to participant responses. For justifiability, we asked them to "imagine that you are considering the purchase of an expensive piece of clothing, jewelry, home goods, or sports equipment." We then asked them to state their level of agreement with statements regarding justifiability (labeled J#) shown in the Appendix. They were then told, "now consider that during a promotion at work you were to earn, as an award, the item that you considered purchasing." Participants were then asked to state their level of agreement with the statements about social reinforcement (labeled S#). Participants were asked to respond to all statements on a five point Likert type scale where 1 = strongly disagree and 5 = strongly agree.

Results and Discussion

We gathered three-hundred and four responses for study 1. Fifty-four percent of the respondents were male and 46% were female. The median age was 33 and median income was in the range of $30,000 to $40,000. While slightly below the $41,000 median income estimated

by the Census Bureau, it is most likely an artifact of the pool of potential respondents, as participants recruited from MTurk tend to be younger than the population as a whole and have a lower income ("Understanding Amazon Mechanical Turk," 2015). The authors feel that this difference does not discredit the results of this survey.

Since the intent of this study was to develop a scale to measure employee perceptions of tangible incentives, we performed two screens of the data. First, participants who had never been eligible to receive a tangible incentive (22) were removed from the data prior to analysis. An additional ten respondents reported working zero hours per week and therefore they were also removed prior to analysis as they were not working adults. One respondent was removed for both reasons. This led us to remove a total of 31 responses, leaving a final sample of 273. Even though we removed 31 responses from the analysis, all people who completed the survey were paid one dollar for their participation.

We explored discriminant and convergent validity using the Spearman rank correlations of the responses as shown in Table 1. The correlations within responses to justifiability statements (J#) are significant, as are the correlations within responses to statements regarding social reinforcement (S#). This result suggetsts convergent validity. Divergent validity is suggested by a low level of correlation between the statements on justifiability and the statements about social reinforcement. These results suggest that the statements are measuring what they are meant to measure and not the other construct (Trochim, 2000).

Table 1

Spearman Correlations – Study 1

Statement	J1	J2	J3	J4	J5	J6	S1	S2	S3	S4	S5
J1	1.00	0.56	0.53	0.43	0.67	0.49	0.10	0.04	0.06	0.10	0.10
J2	0.56	1.00	0.39	0.25	0.45	0.35	0.09	0.16	0.13	0.06	0.15
J3	0.53	0.39	1.00	0.47	0.53	0.67	0.06	0.03	0.10	0.06	0.04
J4	0.43	0.25	0.46	1.00	0.52	0.45	0.05	0.06	0.08	0.05	0.22
J5	0.67	0.45	0.53	0.52	1.00	0.57	0.16	0.08	0.14	0.10	0.17
J6	0.49	0.35	0.67	0.45	0.57	1.00	0.04	-0.01	0.00	-0.01	0.08
S1	0.10	0.09	0.06	0.05	0.16	0.04	1.00	0.58	0.49	0.73	0.44
S2	0.04	0.16	0.03	0.06	0.08	-0.01	0.58	1.00	0.59	0.62	0.39
S3	0.06	0.13	0.10	0.08	0.14	0.00	0.49	0.59	1.00	0.58	0.49
S4	0.10	0.06	0.06	0.05	0.10	-0.01	0.73	0.62	0.58	1.00	0.44
S5	0.10	0.15	0.04	0.22	0.17	0.08	0.44	0.39	0.49	0.44	1.00

Correlations > .12 significant at $p < .05$, Correlations > .16 significant at $p < .01$, Correlations > .22 significant at $p < .001$

We analyzed responses to all statements using exploratory factor analysis in MPLUS 7.1. Using an oblique Geomin rotation, we explored solutions from one to four factors. Fit statistics from this analysis can be seen in Table 2. We rejected the one factor solution due to poor fit statistics. We rejected the 4-factor solution, as the improvement from the 3-factor solution was poor. The discriminant validity of the 3-factor solution is weak, as the correlation between factors 1 and 3 was 0.50. In addition, we evaluated the scree plot where there was a large drop in eigenvalues after the 2nd factor. These results caused us to reject the 3 factor solution and use the 2-factor solution. One potential concern regarding the two factor model is the 0.12 correlation ($p < .05$) between the two rotated factors. This raises some concern about discriminant validity which we examine further after Study 2.

Table 2

EFA resuts – Study 1

MODEL	Chi-Square (DF)	CFI	TLI	RMSEA	RMSR
1 FACTOR	754.9 (44)	0.45	0.32	0.24	0.24
2 FACTORS	143.6 (34)	0.92	0.86	0.11	0.04
3 FACTORS	70.1 (25)	0.96	0.92	0.08	0.02
4 FACTORS	33.7 (17)	0.99	0.96	0.06	0.03

Our next step was to figure out which components should remain in the final data model. We first performed an exploratory factor analysis with the responses to all 11 statements. Based upon acceptable loadings for each statement in the EFA, we performed a confirmatory factor analysis using responses to all 11 statements and two factors in order to examine both the fit statistics and the component loadings. The results from both of these analyses are shown in Table 3. For the justifiability factor, the most concerning factor loading was 0.58, coming from statement J4, regarding whether or not people would consider the purchase of the item as splurging. There was also a potential problem with a 0.59 loading for statement J2, regarding whether or not the person would have difficulties justifying the purchase to others. For social reinforcement, the most problematic loading of 0.58 came from statement S5, which asked whether other people would be likely to ask the individual about the reward. Although low, all are above the recommended cutoff level of 0.4 (Fornell & Larcker, 1981). Nevertheless, we reworded some of these statements for Study 2.

To test whether these statements should remain in the scales, we performed an additional CFA without those variables. While the factor loadings of the remaining statements improved, the fit statistics were negatively affected. CFI dropped from 0.91 to 0.70 and the TLI dropped

from 0.89 to 0.64. The RMSEA also rose from 0.10 to 0.17 and the SRMR went from 0.05 to 0.19. Based upon the poorer fit statistics of the revised model, we decided to retain all responses for further analyses.

Table 3

Factor Analyses – Study 1

	Exploratory Analysis		Confirmatory Analysis	
	Factor 1 (Justifiability)	Factor 2 (Social Utility)	Factor 1 (Justifiability)	Factor 2 (Social Utility)
J1	**0.76**	0.07	0.77	
J2	**0.58**	0.10	0.59	
J3	**0.71**	0.02	0.70	
J4	**0.58**	0.01	0.58	
J5	**0.80**	0.11	0.81	
J6	**0.72**	-0.05	0.70	
S1	0.06	**0.79**		0.79
S2	0.00	**0.77**		0.78
S3	0.04	**0.72**		0.72
S4	0.00	**0.86**		0.87
S5	0.10	**0.57**		0.58

Geomin Rotated Factors

Both scales were highly reliable, with the justifiability scale having a composite reliability (CR) of 0.85 and Chronbach's alpha of 0.76, and the social reinforcement scale having CR = 0.86 and $\alpha = 0.83$. As alluded to earlier, we had potential issues with discriminant valididty. Confirming this, the average variance explained (AVE) was low for both scales, with the justifiability scale having an AVE of only 0.48 and the social reinforcement scale having an AVE of 0.56. While the latter is above the traditional cutoff of 0.50, the AVE for the justifiability scale was below this level. Because of these results, we made some modifications to the statements for Study 2.

The results from Study 1 show that these two latent variables exist and are sufficiently well measured by responses to the existing statements. The low levels of factor loadings on two statements on the justifiability scale were troubling, as was the low level of average variance explained. Removing them did improve the AVE but the other fit statistics deteriorated. Thus, while promising, we do not feel that the set of statements, as worded, adequately capture the latent variables. This led us to run a modified set of statements for Study 2.

Study 2

Given some of the problematic statistics found in Study 1, we modified the wording of one statement on the justifiability scale and added an additional statement. The statement we changed was "I would consider the purchase of such an item splurging." Based upon a closer review of construct definition in Jeffrey and Shaffer (2007), we rephrased the statement as: "If I were to buy this with my own money, it would be an extravagant purchase." We felt this statement better addressed the construct of justifiability. Based upon our review of the justifiability construct, we added an additional statement, "buying this with my own money would be self-indulgent."

Although there were some loading issues, we chose not to modify the wording of the second justifiability statement (regarding justifying a purchase to others). This decision was made since this statement was worded almost identically to the first, which had no loading problems. In addition, this statement addresses one of the central ideas of justifiability, the need to justify purchases to another person. The social reinforcement statement with poor loading was "people will be likely to ask me how I earned this award." Upon review of the original construct, the authors felt that this did not directly address the theoretical construct. The original research claimed that the performance leading to the award would be a topic brought up by others. We therefore reworded this statement to be "People will ask me about the performance that led to the earning of this award."

Results and Discussion

One-hundred and ninety-eight responses were gathered for study 2. As with Study 1, we removed respondents who had never been eligible to receive tangibles and those who do not work. This left us with a usable sample of 180 respondents. Fifty-three percent of our respondents were male, 46% were female, and one respondent replied "other".

We followed the same analysis steps we used in Study 1. We first examined discriminant and convergent validity using the Spearman rank correlations of the statements as shown in Table 4. As with study one, convergent validity was demonstrated by the high levels of correlation within the respective constructs and low levels of correlation across different construct statements. This suggests that the statements are measuring what they are meant to measure and not another construct (Trochim, 2000). The only problematic correlation was between two new statements. Responses to the statement "People will ask me about the performance that led to the earning of this award" (S5) and "If I were to buy this with my own money, it would be an extravagant purchase" (J4), were correlated at $r = .24$, $p < 0.01$. This correlation was also present in study 1 before the re-wording of the statements. This leads us to

believe that there are some commonalities in the psychological mechanisms that lead to the responses to these statements. This will be discussed at the end of the article.

Table 4

Statement	J1	J2	J3	J4	J5	J6	J7	S1	S2	S3	S4	S5
J1	1.00	0.55	0.58	0.48	0.61	0.51	0.47	-0.03	-0.03	-0.09	-0.03	0.00
J2	0.55	1.00	0.43	0.34	0.53	0.33	0.31	0.06	0.00	-0.10	-0.04	-0.02
J3	0.58	0.43	1.00	0.56	0.54	0.67	0.64	-0.07	-0.01	-0.01	0.00	0.15
J4	0.48	0.34	0.56	1.00	0.51	0.49	0.66	0.04	0.09	0.10	0.11	0.24
J5	0.61	0.53	0.54	0.51	1.00	0.54	0.49	0.06	0.03	-0.02	0.07	0.11
J6	0.51	0.33	0.67	0.49	0.54	1.00	0.65	-0.12	0.00	0.01	-0.04	0.14
J7	0.47	0.31	0.64	0.66	0.49	0.65	1.00	-0.02	0.11	0.08	0.08	0.25
S1	-0.03	0.06	-0.07	0.04	0.06	-0.12	-0.02	1.00	0.58	0.53	0.68	0.49
S2	-0.03	0.00	-0.01	0.09	0.03	0.00	0.11	0.58	1.00	0.57	0.68	0.44
S3	-0.09	-0.10	-0.01	0.10	-0.02	0.01	0.08	0.53	0.57	1.00	0.60	0.58
S4	-0.03	-0.04	0.00	0.11	0.07	-0.04	0.08	0.68	0.68	0.60	1.00	0.53
S5	0.00	-0.02	0.15	0.24	0.11	0.14	0.25	0.49	0.44	0.58	0.53	1.00

Spearman Correlations – Study 2 Correlations > .14 significant at $p < .05$, Correlations > .18 significant at $p < .01$, Correlations > .25 significant at $p < .00$

We next analyzed responses to all statements using exploratory factor analysis in MPLUS 7.1, exploring solutions from one to four factors. Output from this analysis is in Table 5. We rejected the one factor solution due to low fit statistics and the 4-factor solution did not converge. This left us two potential solutions to examine. The discrimant validity of the 3-factor solution was poor, as the correlations between the factors were significant, $\rho(1,2) = -0.18$, $p < .05$, $\rho(2,3) = -0.13$, $p < .05$. In addition, the eigenvalue for factor 3 is only 1.08, compared to 4.10 and 3.28 for factors 1 and 2 respectively. Finally, the chi-squared value of the model did not attain significance, $\chi^2(33) = 41.81$, $p = 0.14$. Based upon these statistics, we used the 2-factor solution for future analysis.

Table 5

EFA results – Study 2

MODEL	Chi-Square (DF)	CFI	TLI	RMSEA	RMSR
1 FACTOR	563.62 (54)	0.51	0.40	0.23	0.20
2 FACTORS	131.56 (43)	0.91	0.87	0.11	0.05
3 FACTORS	41.81 (33)	0.99	0.98	0.04	0.02
4 FACTORS	Did Not Converge				

Our next step was to figure out the components that should remain in the final data model, loading each statement on the factor for which the loading was stronger in the EFA shown in table 6. We performed a confirmatory factor analysis using all 12 statements and examined both the fit statistics and the factor loadings. For the justifiability factor, the most concerning factor loading came from the statement regarding whether or not the person would have difficulties justifying a purchase to others. This statement loaded at only 0.55, lower than the 0.59 found in Study 1. Given that we believed this concept as central to the construct of justifiabiltity, we examined the responses further.

Table 6

Factor Analyses – Study 2

	Exploratory Analysis		Confirmatory Analysis	
	Factor 1 (Justifiability)	Factor 2 (Social Utility)	Factor 1 (Justifiability)	Factor 2 (Social Utility)
J1	**0.73**	-0.03	0.73	
J2	**0.63**	-0.05	0.54	
J3	**0.82**	-0.01	0.82	
J4	**0.70**	0.16	0.70	
J5	**0.70**	0.05	0.71	
J6	**0.71**	-0.03	0.70	
J7	**0.76**	0.12	0.75	
S1	-0.05	**0.77**		0.78
S2	0.01	**0.80**		0.80
S3	0.00	**0.70**		0.70
S4	0.00	**0.88**		0.88
S5	0.15	**0.62**		0.62

We believed that responses to this statement may differ by marital status, but ANOVA analysis showed no significant differences by marital status, $F(169,3) = 1.28$, $p = 0.28$. Comparing the individual answers using a Tukey Honestly Significant Difference test, the only comparison approaching significance was between previously married people and those living with a partner.

Because of these disappointing results, we felt that responses to this statement may not adequately address the concept of justifiability. Therefore, we removed this statement and performed the CFA analysis again. The removal of the responses to this statement did not improve the fit of the model so we kept it for our future analyses. Our modified statement S5 (people asking about performance) also had poor loadings (0.62), but it was better than the result in study one, where the original statement loaded at 0.58.

Both revised scales were highly reliable, with the justifiability scale having a composite reliability (CR) of 0.88 and Chronbach's alpha of 0.87, and the social reinforcement scale having CR = 0.87 and $\alpha = 0.87$. The average variance explained (AVE) was 0.51 for the justifiability scale and 0.58 for the social reinforcement scale, both above the traditional cutoff of 0.50 (Fornell & Larcker, 1981). These results suggest that our scales have good validity.

To further examine the validity of these scales, we examined how they correlated with the demographic variables with which they should correlate (i.e. predictive validity). We created two latent variables by averaging responses to the final statements used for the scales as outlined. As predicted, justifiability decreased with income, with a .16 decrease in the average justifiability score for each $10,000 of additional reported income, $t(178) = 2.09$, $p < .05$. As income rises, justification of a purchase should be easier to justify, confirming the psychology behind the justifiability construct.

We believed that social utility would be higher for individuals who had a "significant other" who would provide said utility. To test this belief, we performed an ANOVA for the scores on the social reinforcement variable and found that scores did differ depending on marital status, $F(3, 176) = 4.0$, $p < .01$. Using a Tukey HSD test, participants who were currently married reported higher levels of social utility than those who had never been married, $M_{married} = 3.74$, $SE = .13$. $M_{never\ married} = 3.29$, $p < .05$. Those reporting living with a partner reported a higher level of social utility that those who had never been married, although the difference only approached significance, $M_{living\ with\ partner} = 3.75$, $SE = .63$, $p = .07$. These results show good predictive validity of this scale.

Based upon these analyses, we find sufficient support to use the revised scales for measuring justifiability and social reinforcement. The fit statistics for the overall model as well as the factor loadings improved in Study 2. In addition, we find that the scores vary with the demographic variables with which they should vary. While we refrain from calling the fit perfect, we feel it is a good way to access these latent variables.

General Discussion and Conclusion

The purpose of this research was to create scales to measure the constructs of justifiability and social reinforcement. In two studies, we finalized a set of six statements for justifiability and five statements for social reinforcement The fit statistics for these scales all exceed the recommended cutoffs (Fornell & Larcker, 1981), as do all of the factor loadings for the individual statements. While it is possible that there are ways to word these statements that would improve the overall fit, we find that the fit is sufficient to move forward with these statements.

Limitations and Future Research

One limitation of study is that it only measured two of the motivating properties of tangible incentives. This might contribute to potential problems of statistical power. Future research should add an analysis of the additional factors proposed in the literature. A related limitation is the small pool of initial statements analyzed. A small starting pool did not allow the authors as much lattitude to identify and remove poorly performing statements.

We have ongoing concerns regarding about the apparent need to drop statement J2 regarding justifying a purchases to others. The authors believe that this is one of the fundamental elements of justifiability. The factor loading on this was low, even though it was worded almost identically to the statement regarding the need to justify a purchase to oneself. Future research should examine better ways to address this psychological process.

We also have a lingering concern about the correlation of two items in different scales. Statements J4 (if I were to buy this with my own money, it would be an extravagant purchase) and S5 (people will ask me about the performance that led to the earning of this award) are significantly correlated in both studies. While the correlations are small, .22 in study 1 and .24 in study 2, they are still concerning. It is possible that people think that others are more likely to ask about extravagant purchases. This is clearly speculation but future work should try to identify the underlying psychological mechanisms in order to better identify the underlying constructs of justifiability and social reinforcement. This will likely include tests of different wording of some of the statements used.

Theoretical and practical implications

We now have two valid scales that can be used to assess justifiability and social utility. Research can now assess these constructs to determine the extent to which they contribute to the motivational power of tangible incentives. This will move us beyond aggregate level studies that show the primacy of cash or non-cash incentives and begin to answer questions regarding individual differences in motivational power. We hope that researchers will use these scales when comparing the performance uplift of different types of incentives.

In addition, firms can gather employee resonses to these statements and use these responses to understand how effective different types of incentives (e.g. cash vs. non-cash) will be with different employees. Rather than looking only at mean performance improvements, these employee responses will allow firms to better individualize incentive plans.

Conclusion

The goal of this research was to develop scales for the latent variables of justifiability and social reinforcement. While not perfect, the scales developed in this paper are a good first step at discovering ways to measure these constructs that make tangible incentives motivating. We encourage researchers to utilize these scales in order to add additional validity assessments. We also encourage researchers to experiment with the wording of these statements and potentially add new ones that will better address these latent constructs.

References

Alport, G. W. (1954). The historical background of modern psychology. In G. Lindzey (Ed.). Cambridge, MA: Adison-Wesley.

Bandura, A. (1969). *Principles of behavior modification*. New York: Holt, Rinehart, & Winston.

DeVellis, R. F. (2012). *Scale Development: Theory and Applications*. SAGE Publications.

Fornell, C., & Larcker, D. F. (1981). Structural Equation Models With Unobservable Variables and Measurement Error: Algebra and Statistics. *Journal of Marketing Research (JMR), 18*(3), 382–388.

Hsee, C. K. (1996). Elastic justification: How unjustifiable factors influence judgments. *Organizational Behavior and Human Decision Processes*, *66*, 122–129.

Incentive Marketplace Estimate Research Study | Research | The Incentive Research Foundation. (2016, July 21). Retrieved June 16, 2017, from http://theirf.org/research/incentive-marketplace-estimate-research-study/1836/

Jeffrey, S. A. (2003). *Non-Monetary incentives and motivation: When is Hawaii better than cash?*

Jeffrey, S. A. (2009). Justifiability and the Motivational Power of Tangible Noncash Incentives. *Human Performance, 22*(2), 143–155.

Jeffrey, S. A., & Adomdza, G. K. (2011). Incentive salience and improved performance. *Human Performance*, *24*(1), 47–59.

Jeffrey, S. A., & Shaffer, V. (2007). The motivational properties of tangible incentives. *Compensation and Benefits Review, 39*(3), 44–44.

Luthans, F., & Stajkovic, A. D. (2000). Provide recognition for performance improvement. In E. A. Locke (Ed.), *Principles of Organizational Behavior* (pp. 166–180). Oxford, England: Blackwell.

Mahoney, M. J. (1974). *Cognition and behavior modification*. Cambridge, MA: Ballinger.

Scopelliti, I., Loewenstein, G., & Vosgerau, J. (2015). You call it "self-exuberance"; I call it "bragging": Miscalibrated predictions of emotional responses to self-promotion. *Psychological Science, 26*(6), 903–914. https://doi.org/10.1177/0956797615573516

Shaffer, V. A., & Arkes, H. R. (2009). Preference reversals in evaluations of cash versus non-cash incentives. *Journal of Economic Psychology*, *30*(6), 859–872.
Stajkovic, A. D., & Luthans, F. (1997). A meta-analysis of the effects of organizational behavior modification on task performance, 1975-1995. *Academy of Management Journal*, *40*(5), 1122–1149.
Trochim, W. (2000). *The Research Methods Knowledge Base*. Cincinatti, OH: Atomic Dog.
Understanding Amazon Mechanical Turk. (2015, November 5). Retrieved August 9, 2017, from http://consumerresearch.georgetown.edu/research/understanding-amazon-mechanical002Dturk/
Wright, T. A., Quick, J. C., Hannah, S. T., & Blake Hargrove, M. (2017). Best practice recommendations for scale construction in organizational research: The development and initial validation of the Character Strength Inventory (CSI). *Journal of Organizational Behavior*, *38*(5), 615–628.

Appendix

Scale Statements

	Study 1	Study 2
J1	I would have a hard time justifying a purchase like this **to myself** if I had to use my own money	I would have a hard time justifying a purchase like this **to myself** if I had to use my own money
J2	I would have a hard time justifying a purchase like this **to someone else** if I had to use my own money	I would have a hard time justifying a purchase like this **to someone else** if I had to use my own money
J3	There are many better things I could purchase with the amount of money this would require.	There are many better things I could purchase with the amount of money this would require.
J4	I would consider the purchase of such an item splurging.	If I were to buy this with my own money, it would be an extravagant purchase.
J5	I would feel guilty if I purchased this item.	I would feel guilty if I purchased this item.
J6	There are more justifiable ways to spend the amount of money this would cost.	There are more justifiable ways to spend the amount of money this would cost.
J7		Buying this with my own money would be self-indulgent.
S1	I would talk with many people about this award	I would talk with many people about this award
S2	Knowing that others know that I earned this award would make me happy	Knowing that others know that I earned this award would make me happy
S3	People will admire the award I earned.	People will admire the award I earned.
S4	I would enjoy talking to people about this award.	I would enjoy talking to people about this award.
S5	People will be likely to ask me how I earned this award.	People will ask me about the performance that led to the earning of this award.

