Journal of Economics and Behavioral Studies (ISSN: 2220-6140)
Vol. 13, No. 2, pp. 22-33, April 2021

Meta-Analysis of Seven Standard Experimental Paradigms Comparing Student to Non-student Populations

Jorge N. Zumaeta
Florida International University, USA
jzumaeta@fiu.edu

Abstract: This study embarked on the very challenging proposition of systematically organizing and classifying an assortment of experimental economics essays pertinent to seven experiments performed with both non-student and student populations. The experiments were the Dictator game, Stag Hunt – Coordination game, Risk Aversion Measurement (as measured by the players type of lottery choice), Trust game, Guessing game, Prudence Measurement, and the Guessing game. This meta-analysis reviewed 126 published and unpublished papers collected from several journals and papers provided by several authors via the Google Groups "Economic Science Association - Experimental Methods Discussion" group. Ultimately, only 39 studies were utilized due to methodological alignment. While some studies showed statistically significant differences between non-students and students as indicated by their respective 95% confidence intervals, the overall random-effects model of each of the seven games showed not to be statistically significant. This study contributes to the literature in three important ways. First, the study generates a comprehensive inventory and review of experiments comparing student to non-student populations for the last four decades; second, the study points out a possible limitation when combining several studies of the same game, despite following similar protocols, suggesting that compounded contextual complexities might diminish aggregate effects of the individuals' behavioral responses to the financial incentives, and third, the study indicates that generalizations from one experimental economic study, may not render a solid base for extending statistical extrapolations applicable to the total population since the aggregate effects do not indicate substantial differences.

Keywords: *Experimental Economics, Economic Attitudes, Financial Decisions, Welfare, Game Theory.*

1. Introduction

A meta-analysis is an explanatory method utilized to summarize quantitative data derived from numerous distinctive research inquiries. The main goal of conducting a meta-analysis is to measure the degree of effect of the exogenous variable on the endogenous variable. Generally speaking, a meta-analysis employs statistical findings from each independent research as the units of estimation in all-inclusive research. These findings are aggregated by effect size or mean differences to obtain an estimated size of the effect (Becker, 1986). The contribution of the meta-analysis pivots on the fact that the indices of effect magnitude are scale-free. The scale-free index of an effect magnitude makes it possible to combine the findings of studies that follow the same guidelines of implementation. The meta-analysis was performed utilizing the most recent version of Medcalc form 18.9 (http://www.medcalc.org; 2018). In this study, the effects of group type (students versus non-students) on outcomes of each of the seven different experimental tasks, as indicated by the results of individual studies, were examined to determine the size of the overall effect. Heterogeneity between studies is assessed by Q and I^2 tests (Li et al., 2015). Q is considered a conventional measure of heterogeneity, called Cochran's Q, which is calculated as the weighted sum of squared differences between each study effect and the pooled effect across studies, with the weights being those used in the pooling method. The I^2 statistic describes the percentage of disparity among research studies that is due to heterogeneity rather than chance[1].

[1]In meta-analysis, the regular way of evaluating whether a set of single research studies are homogeneous is by means of the Q test. The Q test let us know about the existence of heterogeneity, but it does not inform us on the degree of such heterogeneity. In the last few years, the I^2 index has been interpret as measuring the degree of heterogeneity in a meta-analysis. In this paper, the Q test and the confidence interval around the I^2 index of all collected studies are compared by using the mean of the Monte Carlo simulation. The calculation indicates the utility of the I^2 index complementing the information provided by the Q test (Huedo-Medina Et all, 2006).

Heterogeneity is defined as significant when $p < 0.05$ in Q statistics and $I^2 > 40\%$ (Graziani, 2014). The results of the meta-analysis will be presented as a weighted mean difference and 95% confidence interval, with a significance level $\alpha = 0.05$. The data of each study and meta-analysis will be summarized with Forest plots. Additionally, due to the multiple types of estimates reported in each study, it was necessary to perform a generic inverse-variance approach to the meta-analysis. The inverse variance method focuses on the weight given to each study that is chosen to be the inverse of the variance of the effect estimate (i.e., one over the square of its standard error). In other words, the larger the research studies with smaller standard errors, are allocated more weight than smaller research studies, which have larger standard errors. It is well known that the student population is convenient due to relatively easy access by professors and flexible schedules while the general population may experience long commuting distances and rigid working schedules. Conducting this meta-analysis was very important due to the strong dependence on the student population recruited for social experimentation by economists and other social scientists. It is important to illustrate that scientific generalizations derived from experimental economic games may lead to challenges when attempting to develop a better understanding of the general population's financial decision-making and behavioral responses.

2. Criteria for Study Selection (Inclusion Criteria)

The inventory and analysis of papers to be considered for this study followed a rigorous and comprehensive selection process that is described in this section. As per Hedges and Becker (1986), there are three main requirements for any paper to be able to be included as a part of a meta-analysis. The first requirement is for the studies' methodology and calculations to follow similar guidelines, the second requirement asks for the studies to be independent, and the third condition requires that the studies must measure the same statistical parameters. Furthermore, the researchers sought additional characteristics, first, the researcher requested that all studies follow the same framing and game design; second, for all studies to compare students to non-student populations; Third, all studies must have included the specific dependent variables to be measured for each of the seven games. Fourth, all studies had to have adequate statistical data that would allow for the calculation of the effect size. For those studies that conducted descriptive analyses, necessary statistics for the transformations into effect sizes would include means and standard deviations. For studies that used inferential statistics, the appropriate transformation equations were applied, such as the conversion of regression coeffects to t statistics by utilizing its standard error.

3. Methodology

Experimental studies in economics are traditionally conducted using undergraduate students as subjects. A longstanding methodological question has been to what extent the results reported from this subset of the population generalize to different societal groups. This meta-analysis reviewed 126 published and unpublished papers collected through a comprehensive identification process from several journals and papers provided by several authors via the Google Groups "Economic Science Association - Experimental Methods Discussion" group. Ultimately, the list was narrowed down to only 39 studies due to methodological alignment in terms of utilizing representative samples of student versus non-student populations, same experimental economics' protocols (number of experiment rounds), and reasonable incentives to participate (financial compensation). The sample size varies from study to study; nevertheless, the study only included studies with a random sample selection rather than samples of convenience. The 39 studies represent all studies available that meet the minimum requirements to be included in this meta-analysis.

Dictator Game: The first game to be examined is the Dictator game. In the dictator game, the first player, "the dictator", determines how to split an endowment (such as a cash prize) between themselves and the second player (Andreoni et al., 2008). The dictator's action space is complete and therefore is at their own will to determine the division, which means that the recipient has no influence over the outcome of the game (Engel, 2011). In the meta-analysis, seven studies compared the observed donations of students to non-students. Table 1 below shows the results of the seven studies which compared students with non-students. Bekkers (2007) conducted a field study using a modification of the dictator game in a large random sample survey in the Netherlands (n=1,964). The amount of money earned with survey participation was included as a measure of the stakes in the donation decision.

Stakes varied from $7.5 to $13.75 and were $11 on average. Beramendi (2016) conducted lab experiments at the experimental lab with 66 University of Oxford students from the subject pool. Subjects were asked to allocate an endowment of 1000 ECUs between them and another randomly selected participant in the room. Participants were informed that only half of them would receive the endowment, and the ones who received the endowment would be randomly paired with those who didn't. In a study conducted by Cappelen et al. (2014, 2015), they compared the lab behavior of a student group and a non-student group, where the non-student group on all observable factors was almost identical to the representative adult population in Norway. Carpenter (2005) conducted a study co in field settings with naturally occurring variation in "social framing." The study compared the choices of subject managers drawn from a standard undergraduate population with subject managers drawn from the executive MBA (EMBA) program at Case's Weatherhead School of Management. In a study conducted by Charness et al. (2007), they explored the effects of social distance in experiments conducted over the Internet on three continents, in-classroom laboratory sessions conducted in Israel and Spain, and in computer sessions pairing participants from different states.

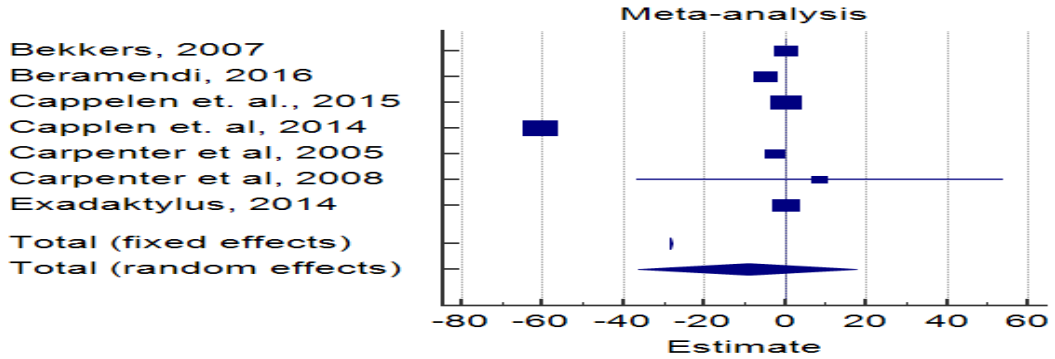
Participants were students at Middlebury College, nontraditional students at Kansas City Kansas Community College (KCKCC), and employees at a Kansas City distribution center. The stake in the dictator game was $100. Exadaktylos (2013) conducted a study in Spain with A total of 835 individuals participating in the experiment. One out of ten participants was randomly selected to be paid. The average earnings among winners, including those winning nothing (18.75%), were €9.60. Bekkers (2007), Beramendi (2016), Cappelen et al. (2014, 2015), Carpenter (2005), and Exadaktylos (2013) all found significant differences in mean donations between students and non-students 95% confidence intervals for the difference (non-students – students) estimates are reported in Table 1 below. Non-students on average gave significantly more donations than students in each study. However, for donations, the total random effect model was not statistically significant, t = -9.972, SE = 13.977, 95% CI [-37.367, 17.422], p =0.476. A random-effect model was chosen because of evidence of heterogeneity based on a significant test for heterogeneity, Q = 7278014.82, p < .001, I^2 = 100.00%. The Forest plot can be found below (figure 1). The meta-analysis found no overall statistically significant differences in the amount of donations given by students and non-students.

Table 1: Meta-Analysis Results for Dictator Game Comparing Non-Students to Students

Study	Estimate*	Standard Error	95% CI	z	P	Weight (%)		N
						Fixed	Random	
Bekkers, 2007	0.35	0.04	0.278 to 0.424			8.94	14.90	1964
Beramendi, 2016	-4.73	0.04	-4.809 to -4.648			7.31	14.90	585
Cappelen et al. 2015	2.27	0.02	2.230 to 2.316			25.27	14.90	354
Capplen et al. 2014	-60.47	0.02	-60.500 to -60.436			46.04	14.90	254
Carpenter et al. 2005	-2.73	0.16	-3.043 to -2.424			0.49	14.90	63
Carpenter et al. 2008	0.37	23.12	-44.945 to 45.686			0.00	10.59	488
Exadaktylos, 2013	-1.88	0.03	-1.938 to -1.812			11.95	14.90	765
Total (fixed effects)	-27.82	0.01	-27.839 to -27.796	-2515.124	<0.001	100.00	100.00	4473
Total (random effects)	-9.97	13.98	-37.367 to 17.422	-0.714	0.476	100.00	100.00	4473

* Estimate was based on t statistic from the study.

Figure 1: Forrest Plot of Meta-Analysis for Dictator Game



Stag Hunt – Coordination Game: The second game in the meta-analysis was the Stag Hunt - Coordination game. The stag hunt is a game that describes a conflict between safety and social cooperation. In the meta-analysis, four studies compared the amount of coordination between non-students and students. Cooper (2005) conducted a study that set experiments in a corporate environment where a manager attempts to overcome a history of coordination failure by employees using either financial incentives or communication. Montmarquette et al. (2004), investigated the relationship between executive pay and performance after a merger by dissociating the respective influence of shifts, which occur in both compensation incentives and team composition. They experimented with professional farm apprentice students as well as with generic university students. Ferré et al. (2017), designed an experiment to study incentives for the adoption of more sustainable land management practices using a modified coordination game.

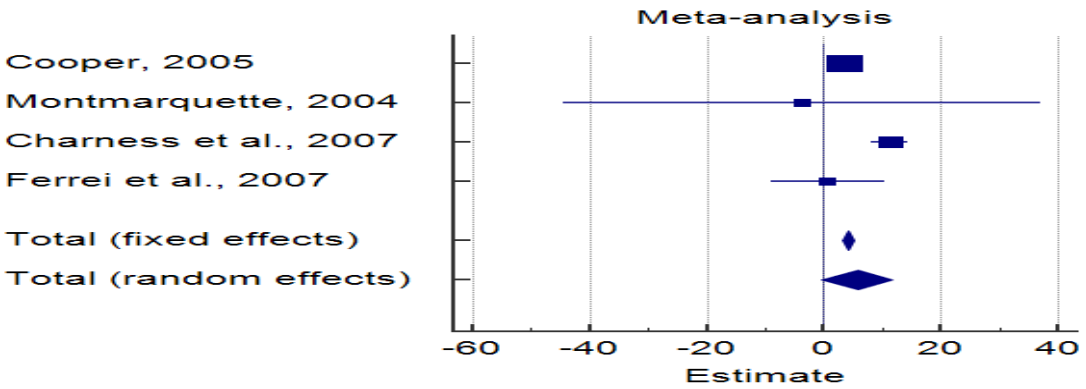
In one treatment, the experiment is highly contextualized and characterizes the situation of farmers and the cultivation of their land, while the other treatment uses abstract and context-free wording. The game was played between students and professionals. Table 1 below shows the results of the four studies. Cooper (2005, 2006) and Charness et al. (2007) found statistically significant differences in coordination between non-students and students with non-students showing more coordination than students as indicated by the positive values of the 95% confidence intervals. For coordination, the total random effect model was not statistically significant, t = 5.658, SE = 3.145, 95% CI [-0.506, 11.822], p = 0.072. A random-effect model was chosen because of evidence of heterogeneity based on a significant test for heterogeneity, Q = 21.522, p = 0001, I^2 = 86.06%. The Forest plot can be found below (figure 2). The meta-analysis found no overall statistically significant differences in coordination between students and non-students.

Table 2: Meta-Analysis Results for Stag Hunt Game Comparing Non-Students to Students

Study	Estimate*	Standard Error	95% CI	Z	P	Weight (%)		N
						Fixed	Random	
Cooper, 2005	3.51	0.53	2.49 to 4.55			89.46	40.51	743
Montmarquette, 2004	-3.77	20.82	-44.58 to 37.04			0.057	2.16	1440
Charness et al. 2007	11.30	1.62	8.13 to 14.48			9.47	36.96	128
Ferré et al. 2017	0.67	4.94	-9.01 to 10.36			1.02	20.37	222
Total (fixed effects)	4.22	0.50	3.25 to 5.20	8.475	<0.001	100.00	100.00	2533
Total (random effects)	5.66	3.15	-0.56 to 11.82	1.799	0.072	100.00	100.00	2533

* Estimate was based on t statistic from the study.

Figure 2: Forrest Plot of Meta-Analysis for Stag Hunt



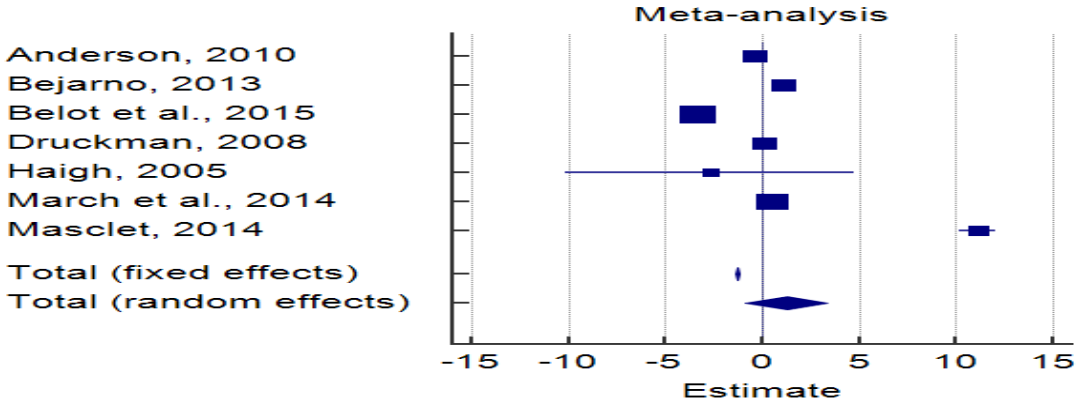
Risk Aversion Measurement: The third game in the meta-analysis was the lottery game which compared risk aversion between students and non-students in seven studies. All studies measured risk aversion by the player type of lottery choice. Table 3 below shows the results of the seven studies. Andsersen (2010), Bejerno (2013), Belot (2015, 2010), March (2014), and Masclet (2009) found statistically significant differences in risk assessment between non-students and students. For risk assessment, the total random effect model was not statistically significant, t = 1.247, SE = 1.085, 95% CI [-0.877, 3.371]. The Trust Game, designed by Berg et al. (1995) and otherwise called "the investment game," is the experiment of choice to measure trust in economic decisions. Anderson (2013) conducted a study that measured trust in samples from three related populations in the upper Midwest of the United States: college students, non-student adults from the community surrounding the college, and adult trainee truckers in a residential training program. A random-effect model was chosen because of evidence of heterogeneity based on a significant test for heterogeneity, Q = 2333.16, p < 0.0001, I^2 = 99.74%. The Forest plot can be found below (figure 3). The meta-analysis found no overall statistically significant differences in risk aversion between students and non-students.

Table 3: Meta-Analysis Results for the Risk Aversion Game Comparing Non-Students to Students

Study	Estimate*	Standard Error	95% CI	Z	P	Weight (%)		N
						Fixed	Random	
Andersen, 2010	-0.40	0.15	-0.69 to -0.10			9.63	15.83	44 254
Bejarano, 2013	1.10	0.17	0.76 to 1.43			7.49	15.81	28
Belot et al. 2015	-3.31	0.07	-3.44 to -3.18			48.29	15.87	200
Druckman, 2008	0.062	0.16	-0.25 to 0.37			8.47	15.82	214
Haigh, 2005	-2.71	3.80	-10.16 to 4.74			0.015	5.38	1062
March et al. 2014	0.52	0.09	0.34 to 0.71			25.06	15.86	3702
Masclet, 2009	11.10	0.46	10.20 to 12.00			1.04	15.44	144
Total (fixed effects)	-1.30	0.05	-1.39 to -1.21	-28.002	<0.001	100.00	100.00	5604
Total (random effects)	1.24	1.08	-0.877 to 3.371	1.151	0.250	100.00	100.00	5604

* Estimate was based on t statistic from the study.

Figure 3: Forrest Plot of Meta-Analysis for Lottery Game



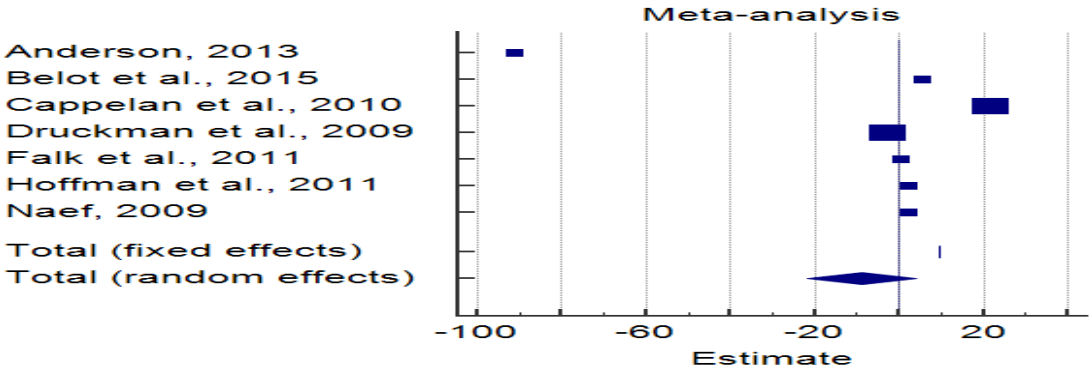
The Trust Game: The fourth game in the meta-analysis was the trust game which measured the amount of trust between students and non-students in seven studies. Belot (2015) in Oxford compared the behavior of students and non-students in the Trust Game. Cappelan (2010) compared student to non-student groups in Norway. Belot (2015) conducted a study in Oxford comparing students with non-students. Druckman (2008) recruited participants from a large, public university in the United States and the surrounding urban community and a substantial number of non-students. Hoffman (2011), conducted a battery of social preference experiments on business people from internet industries as well as on Berkeley students. Table 4 below shows the results of the seven studies. Andserson (2013), Belot (2015), Cappelan (2010), Druckman (2008), Hoffman (2011), and Naef (2009) found statistically significant differences in levels of trust between non-students and students. However, the total random effect model was not statistically significant, t = -8.844, SE = 6.537, 95% CI [--21.657 to 3.970], p = 0.176. A random-effect model was chosen because of evidence of heterogeneity based on a significant test for heterogeneity, $p < 0.0001$, I^2 = 100.00%. The meta-analysis found no overall statistically significant differences in trust between students and non-students. The Forest plot can be found below (figure 4).

Table 4: Meta-Analysis Results for Trust Game Comparing Non-Students to Students

Study	Estimate*	Standard Error	95% CI	z	P	Weight (%)		
						Fixed	Random	N
Anderson, 2013	-91.02	0.32	-91.65 to -90.38			0.0077	14.29	192
Belot et al. 2015	5.40	0.08	5.25 to 5.55			0.14	14.29	128
Cappelan et al. 2010	21.69	0.004	21.68 to 21.70			49.89	14.29	254
Druckman et al. 2008	-2.77	0.004	-2.78 to -2.77			49.89	14.29	200
Falk et al. 2013	0.48	0.61	-0.70 to 1.69			0.0021	14.27	652
Hoffman et al. 2011	2.16	0.11	1.95 to 2.38			0.067	14.29	95
Naef, 2009	2.13	0.57	1.02 to 3.26			0.0025	14.28	193
Total (fixed effects)	9.44	0.003	9.43 to 9.45	3341.505	<0.001	100.00	100.00	1714
Total (random effects)	-8.84	6.54	-21.66 to 3.97	-1.353	0.176	100.00	100.00	1714

* Estimate was based on t statistic from the study.

Figure 4: Forrest Plot of Meta-Analysis for Trust Game



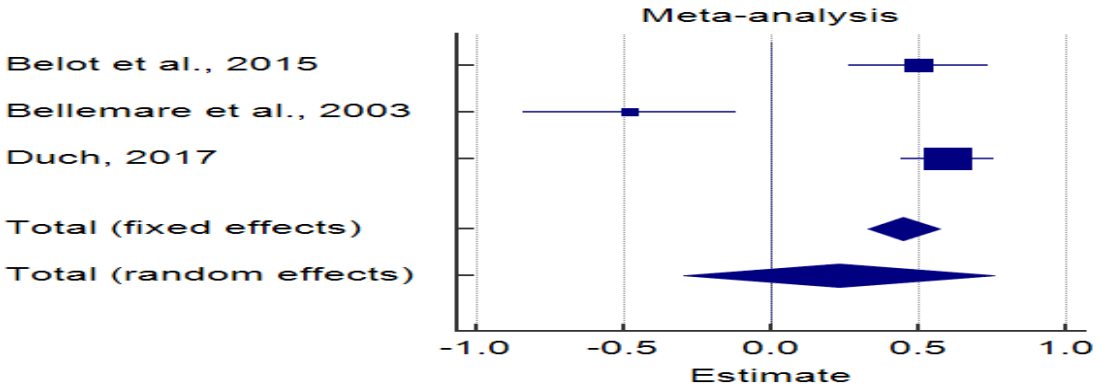
The Guessing Game: The Guessing game (also known as the beauty contest game) was the fifth game to be investigated. In the guessing game, participants are asked to pick a number between 0 and 100, with the winner of the contest being the participant that is closest to 2/3 times the average number picked of all participants. Measured societal trust and trustworthiness by combining the virtues of laboratory experiments and survey data using a representative sample of the Dutch population. Duch (2017) conducted studies in experimental labs in Oxford, Santiago, Chile, and Pune, India to compare the behavior of students and non-students. Table 5 below shows the results of the three studies. Belot, Bellemare, and Duch found statistically significant differences in levels of correct guesses, as defined by the individual guessing the closest to 2/3 times the average, between non-students and students. However, the total random effect model was not statistically significant, t = 0.231, SE = 0.268, 95% CI [-0.293 to 0.756], p =.387. A random-effect model was chosen because of evidence of heterogeneity based on a significant test for heterogeneity Q = 29.014, p < 0.0001, I^2 = 93.11%. The meta-analysis found no overall statistically significant differences in the number of correct guesses between students and non-students. The Forest plot can be found below (figure 5).

Table 5: Meta-Analysis Results for the Guessing Game Comparing Non-Students to Students

Study	Estimate*	Stand. Error	95% CI	z	P	Weight (%)		N
						Fixed	Random	
Belot et al. 2015	0.50	0.12	0.26 to 0.74			27.24	33.86	502
Bellemare et al. 2003	-0.48	0.16	-0.84 to -0.12			11.46	30.96	126
Duch, 2017	0.60	0.08	0.44 to 0.76			61.30	35.19	128
Total (fixed effects)	0.45	0.063	0.33 to 0.57	7.166	<0.001	100.00	100.00	756
Total (random effects)	0.23	0.27	-0.29 to 0.76	0.865	0.387	100.00	100.00	756

* Estimate was based on t statistic from the study.

Figure 5: Forrest Plot of Meta-Analysis for Guessing Game



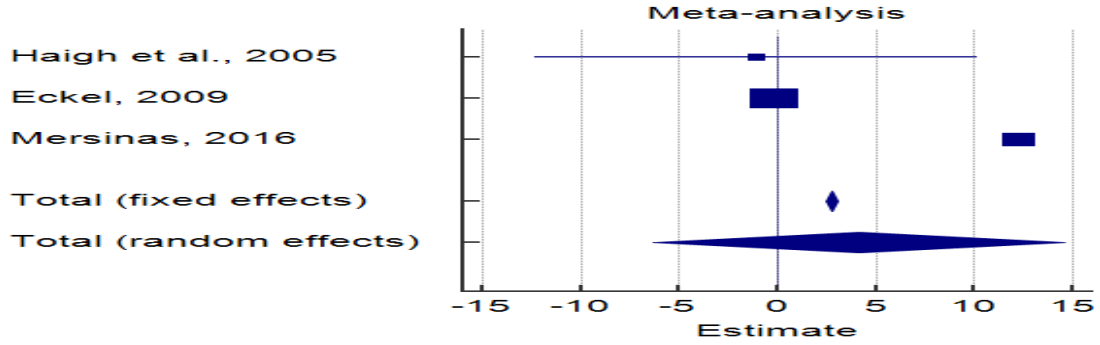
The Prudence task - The Pair-wise Lottery game was the sixth game analyzed in this meta-analysis utilizing three, studies. Haigh (2005) compared both students and non-students, specifically Chicago Board of Trade (CBOT) professional traders and undergraduate students. Grossman (2009) conducted a study in the SCSU Department of Economics Research and Teaching Laboratory. Subjects were recruited by email and posters to participate in a three-part experiment and participation was on a first-come, first-served basis, comparing students and non-students. Mersinas (2016) conducted an experiment with two different samples. The sample of information security professionals was drawn from previous students of the distance learning MSc in Information Security at Royal Holloway, University of London. The student sample was drawn from individuals registered in the database of the Laboratory for Decision Making and Economic Research at RHUL. Table 6 below provides the results of three studies in the analysis. Of the three studies, only Mersinas (2016) found statistically significant findings. However, the total random effect model was not statistically significant, t = 4.163, SE = 5.342, 95% CI [-6.307 to 14.633], p = 0.436. A random-effect model was chosen because of evidence of heterogeneity based on a significant test for heterogeneity Q = 1648.61, p < 0.0001, I^2 = 99.88%. The meta-analysis found no overall statistically significant differences in prudence between students and non-students. The Forest plot can be found below (figure 6).

Table 6: Meta-Analysis Results for the Pair-Wise Lottery Game Comparing Non-Students to Students

Study	Estimate*	Stand Error	95% CI	z	P	Weight (%)		
						Fixed	Random	N
Haigh et al. 2005	-1.06	5.73	-12.29 to 10.17			0.052	25.96	118
Grossman, 2009	-0.22	0.15	-0.51 to 0.074			75.97	37.03	93
Mersinas, 2016	12.21	0.26	11.69 to 12.73			23.98	37.01	118
Total (fixed effects)	2.76	0.13	2.50 to 3.01	21.116	<0.001	100.00	100.00	329
Total (random effects)	4.16	5.34	-6.30 to 14.63	0.779	0.436	100.00	100.00	329

* Estimate was based on t statistic from the study.

Figure 6: Forrest Plot of Meta-Analysis for Pair-Wise Lottery Game



The Ultimatum Game: The Ultimatum game was the seventh game analyzed in this meta-analysis. This meta-analysis combined the results of eight studies which compared minimum accepted offer between students and non-students. Burks (2005) conducted a study in which participants were students at Middlebury College, non-traditional students at Kansas City Kansas Community College (KCKCC), and employees at a Kansas City distribution center. Exadaktylos (2013) conducted a study in which employed data from a survey experiment conducted with a representative sample of a U.S city's population. They reported behavioral data for the ultimatum game which included students and non-students as well as volunteers and non-volunteers. Fu (2007) conducted an experiment on ultimatum games with subjects who were representative of Tawain. The focus was on the size effect of monetary stakes when experimental subjects are "real" people rather than students as in previous studies. These results were compared with the student sample. Güth (2002) conducted a study in which readers of the German weekly, *Die Zeit*, participated in a three-person bargaining experiment. In the data analysis, they focused on student behavior as well as the non-student population.

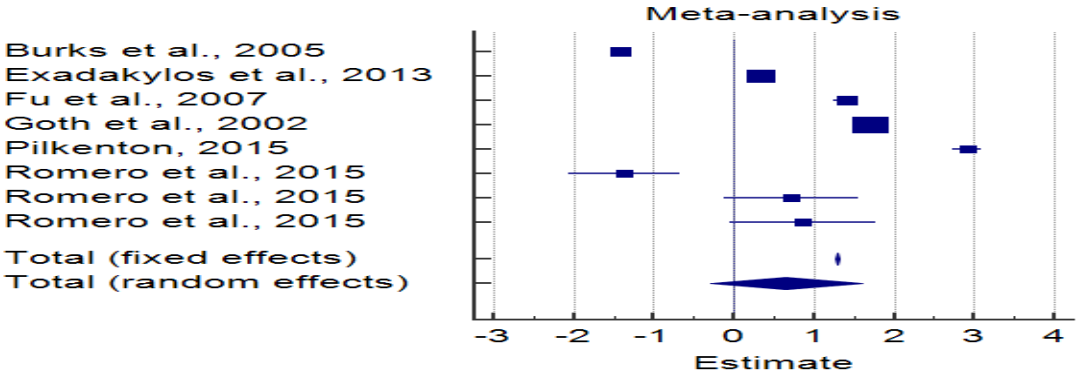
Pilkenton (2015) studied data that was collected over two weeks in two different locations in Davidson, NC– the Davidson Farmers' Market and on-campus at Davidson College. Romero (2015) conducted a study in which the main ethnic groups studied were: Mestizos, Indigenous, Montubios, and African-Ecuadorians. They worked with the standard subject, college students and non-standard subjects, villagers. Table 7 provides a summary of each study in the analysis. Exadaktylos (2013), Fu (2007), Güth (2002), Pilkenton (2015), and Romero (2015) found statistically significant results between non-students and students. Non-students tended to accept a larger minimum offer than students. However, the total random effect model was not statistically significant, t = 0.653, SE = 0.484, 95% CI [-0.295 to 1.602], p = 0.177. A random-effect model was chosen because of evidence of heterogeneity based on a significant test for heterogeneity Q = 9759.66, p < 0.0001, I^2 = 99.93%. %. The meta-analysis found no overall statistically significant differences between students and non-students in terms of accepting a larger number of minimum offers. The Forest plot can be found below (figure 7).

Table 7: Meta-Analysis Results for the Ultimatum Game Comparing Non-Students to Students

Study	Estimate	Stand Error	95% CI	z	P	Weight (%)		
						Fixed	Random	N
Burks et al. 2005	-1.40	0.036	-1.48 to -1.33			5.79	12.94	48
Exadaktylos et al. 2013	0.33	0.021	0.29 to 0.37			17.01	12.94	765
Fu et al. 2007	1.40	0.081	1.23 to 1.56			1.12	12.90	276
Güth et al. 2002	1.70	0.010	1.68 to 1.72			75.02	12.94	995
Pilkenton, 2015	2.91	0.090	2.73 to 3.08			0.93	12.89	105
Romero et al. 2015	-1.37	0.35	-2.07 to -0.68			0.060	12.10	40
Romero et al. 2015	0.72	0.43	-0.12 to 1.56			0.041	11.74	38
Romero et al. 2015	0.85	0.46	-0.059 to 1.77			0.034	11.55	31
Total (fixed effects)	1.29	0.008	1.27 to 1.31	149.269	<0.001	100.00	100.00	2298
Total (random effects)	0.65	0.484	-0.29 to 1.60	1.351	0.177	100.00	100.00	2298

* Estimate was based on t statistic from the study.

Figure 7: Forrest Plot of Meta-Analysis for Ultimatum Game



4. Overall Results and Discussion

As previously stated, the meta-analysis of the 39 studies revealed that while some studies showed statistically significant results between non-students and students as indicated by their respective 95% confidence intervals, the overall random-effects model of each of the seven games showed not to be statistically significant.' With this finding, the study points out a possible limitation of aggregating several studies of each

game, despite following similar protocols, the combined random effect is not significant, suggesting that compounded contextual complexities (specific market conditions during the study, intrinsic cultural factors of the studied population, other nuances of studies, etc.) may cancel or dilute the effects or statistical significance of each study. Furthermore, the research findings suggest that generalizations from one study, may not render a solid base for extending statistical extrapolations applicable to the total population since the aggregate effects do not indicate substantial differences. One motivation for economists for conducting experimental studies is to gauge more accurately how specific populations would respond to financial incentives to economic games. The research utilizing economic games has the practical intention of learning about the decision-making process of specific populations and to potentially serve as the base for generating policies, marketing campaigns, and/or educational interventions.

For example, in 2015, several research studies were conducted on welfare recipients in Miami to learn more about their respective financial decision as a guide to promoting policies supporting self-sufficiency. Putting into context the findings of the meta-analysis of the seven games, we can conclude that the inferences derived from the study on the economic attitudes and financial decisions among welfare recipients in Miami[2], may not be appropriate for other populations around the nation. The Miami study conducted the very same economic games and compared the results to the student population. This specific study did not find any statistical significance between these two groups, meaning that students and non-students (welfare recipients) exhibited similar responses, attitudes, and behaviors when playing the seven games. The researcher of this paper found a great level of support and encouragement from the Google Groups "Economic Science Association - Experimental Methods Discussion" group played a key role in reaching out locally and internationally which rendered a few studies to include in our research. Additionally, the main editor of the Economic Science Association (ESA) journal reached out to its worldwide members to request papers that aligned with our research initiative.

5. Conclusion and Recommendations

The first conclusion of our meta-analysis illustrates that there is not a definitive claim for establishing aggregated correlational behaviors when comparing experimental studies that include students and non-student populations. Having shared this information, we can argue that an important perspective to consider is the need for a larger sample of scientific essays to be included in subsequent meta-studies focusing on this topic. Future meta-analyses should be conducted with a larger number of studies in order to increase the statistical power of the Q test which measures the amount of heterogeneity of the studies. In this way, a more accurate estimate of the true effect size of the combined studies may be obtained. At the time of the collection and analysis of this paper, there was no other research to be incorporated; perhaps, essays on other languages should be considered. The second conclusion of this study arrives is that researchers need to take into account the limitations of combining and aggregating results from different populations at different locations; experiencing different factors at different times, framing and cultural differences, etc (Druckman, 2011 and Laban, 2017).

Although meta-analysis studies could contain flawed data analysis is emerging and passed on to researchers as “updated evidence” (Li et al., 2015). Using this kind of evidence that contains heterogeneous data sets leads to wrong conclusions. Researchers need to be sensitive to the fact that some papers may contain data that did not follow rigorous and comprehensive collection and tabulation processes. Due to this limitation, generalization derived from one specific study following a specific context could not be utilized in the crafting of government policies, marketing campaigns, and/or programmatic interventions. ESA should provide incentives to researchers to engage in further research that includes both students and non-student groups in the hopes of generating a much more comprehensive inventory of studies to conduct a future meta-analysis. Economic Science has become highly dependent on experimental methods to validate and augment principles and theories. It is in the best interest of science that professional economic associations should incentivize further research in this area.

[2] Zumaeta, Jorge N., Economic Attitudes and Financial Decisions Among Welfare Recipients: Rationality and Risk Aversion, March 2020. (Pending Publication).

References

Andreoni, J. H., William, T. & Vesterlund, L. (2008). Altruism in Experiments. The New Palgrave Dictionary of Economics. London: Palgrave Macmillan.

Andersen, S., Harrison, G. W., Lau, M. I. & Rutström, E. E. (2010). Preference heterogeneity in experiments: Comparing the field and laboratory. *Journal of Economic Behavior & Organization*, 73(2), 209-224.

Anderson, J., Burks, S. V., Carpenter, J., Götte, L., Maurer, K., Nosenzo, D., Potter, R., Rocha, K. & Rustichini, A. (2013). Self-selection and variations in the laboratory measurement of other-regarding preferences across subject pools: evidence from one college student and two adult samples. *Experimental Economics*, 16(2), 170-189.

Bejarano, H. D. (2013). Participation, pricing, and perception in markets with externalities.

Bekkers, R. H. (2007). Measuring altruistic behavior in surveys: The all-or-nothing dictator game.

Belot, M., Duch, R. & Miller, L. (2015). A comprehensive comparison of students and non-students in classic experimental games. *Journal of Economic Behavior & Organization*, 113, 26-33.

Belot, M., Duch, R. & Miller, L. (2010). Who should be called to the lab? A comprehensive comparison of students and non-students in classic experimental games. University of Oxford, Nuffield College Discussion Papers (2010-001).

Beramendi, P., Duch, R. M. & Matsuo, A. (2016). Comparing Modes and Samples in Experiments: When Lab Subjects Meet Real People.

Berg, J., John, D., Kevin, M. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 10(123), 122-142.

Cappelen, A. W., Nielsen, U. H., Tungodden, B., Tyran, J. R. & Wengström, E. (2014-2016). Fairness is intuitive. *Experimental Economics,* 19(4), 727-740.

Cappelen, A. W., Nygaard, K., Sorensen, E. & Tungodden, B. (2010). Efficiency, Equality and Reciprocity in Social Preferences: A Comparison of Students and a Representative Population.

Carpenter, J. P., Burks, S. & Verhoogen, E. (2005). Comparing students to workers: The effects of social framing on behavior in distribution games. In Field experiments in economics (pp. 261-289). Emerald Group Publishing Limited.

Carpenter, J., Connolly, C. & Myers, C. K. (2008). Altruistic behavior in a representative dictator experiment. *Experimental Economics,* 11(3), 282-298.

Charness, G., Haruvy, E. & Sonsino, D. (2007). Social distance and reciprocity: An Internet experiment. *Journal of Economic Behavior & Organization,* 63(1), 88-103.

Cooper, D. J. (2005). Are Experienced Managers Experts at Overcoming Coordination Failure?

Cooper, D. J. (2006). Are Experienced Managers Experts at Overcoming Coordination Failure? *The BE Journal of Economic Analysis & Policy,* 5(2), 1-52.

Druckman, J. N. & McDermott, R. (2008). Emotion and the framing of risky choice. *Political Behavior*, 30(3), 297-321.

Druckman, J. N. & Kam, C. D. (2011). Students as experimental participants. *Cambridge handbook of experimental political science,* 1, 41-57.

Duch, R. M., Galvez, T. & Torres, F. (2017). Treatment Effects and Subject Pool Diversity in the Experimental Lab.

Engel, C. (2011). Dictator games: A Meta Study. *Experimental Economics*, 14(4), 583–610.

Exadaktylos, F., Espín, A. M. & Branas-Garza, P. (2013). Experimental subjects are not different. *Scientific reports,* 3(1), 1-6.

Falk, A., Meier, S. & Zehnder, C. (2013). Do lab experiments misrepresent social preferences? The case of self-selected student samples. *Journal of the European Economic Association,* 11(4), 839-852.

Ferré, M., Engel, S. & Gsottbauer, E. (2017). September. External validity of experiments in environmental economics. Framing and subject pool effects among students and professionals. In 19th Annual BIOECON Conference, 21-22 September 2017, Tilburg University, the Netherlands.

Fu, T. T., Kong, W. H. & Yang, C. (2007). Monetary stakes and socioeconomic characteristics in ultimatum games: An experiment with nation-wide representative subjects. Unpublished Manuscript. Available at: http://www. fas. nus. edu. sg/ecs/events/set2007/programme. html.

Graziani, F., Gennai, S., Roldan, S., Discepoli, N., Buti, J., Madianos, P. & Herrera, D. (2014). Efficacy of periodontal plastic procedures in the treatment of multiple gingival recessions. *Journal of Clinical Periodontology*, 41, 63-S76.

Grossman, P. & Eckel, C. (2009). Loving the Longshot: Risk-Taking with Skewed Gambles.

Güth, W., Schmidt, C. & Sutter, M. (2007). Bargaining outside the lab–a newspaper experiment of a three-person ultimatum game. *The Economic Journal*, 117(518), 449-469.

Haigh, M. S. & List, J. A. (2005). Do professional traders exhibit myopic loss aversion? An experimental analysis. *The Journal of Finance*, 60(1), 523-534.

Hedges, L. V. & Becker, B. J. (1986). Statistical methods in the meta-analysis of research on gender differences. In J. S. Hyde, & M. C. Linn (Eds.), The Psychology of Gender: Progress through Meta-analysis Baltimore: The Johns Hopkins University Press.

Hoffman, M. & Morgan, J. (2011). Who's naughty? Who's nice? Social preferences in online industries. University of California, Berkeley.

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F. & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological methods,* 11(2), 193.

Laban P., J. (2017). Culture, Conformity, and Risk Attitudes: An Experimental Analysis. Available at SSRN 3129799.

Li, S., Jiang, H., Yang, H., Chen, W., Peng, J., Sun, M., Zeng, J. (2015). The Dilemma of Heterogeneity Tests in Meta-Analysis: A Challenge from a Simulation Study. *PLOS ONE*, 10(5), e0127538.

Masclet, D., Colombier, N., Denant-Boemont, L. & Loheac, Y. (2009). Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers. *Journal of Economic Behavior & Organization*, 70(3), 470-484.

Mersinas, K., Hartig, B., Martin, K. M. & Seltzer, A. (2016). Are information security professionals expected value maximizers? An experiment and survey-based test. *Journal of Cybersecurity,* 2(1), 57-70.

Montmarquette, C., Rullière, J. L., Villeval, M. C. & Zeiliger, R. (2004). Redesigning teams and incentives in a merger: An experiment with managers and students. *Management Science*, 50(10), 1379-1389.

MedCalc Statistical Software version 18.9. (2018). (MedCalc Software, Ostend, Belgium; http://www.medcalc.org.

Naef, M. & Schupp, J. (2009). Measuring trust: Experiments and surveys in contrast and combination.

Romero, P. P. & Oleas, S. (2015). Bargaining and Ethnicity: A field experiment with students and villagers.

Pilkenton, N. (2015). Racial Differences in Bargaining Behavior: An Ultimatum Game Analysis. *Journal of Economic Behavior and Organization,* 3(4), 367-388.

Zumaeta, N. J. (2020). Economic Attitudes and Financial Decisions among Welfare Recipients: Rationality and Risk Aversion, (Pending Publication).

Zumaeta, N. J. (2021). Specific Preference Parameters among Welfare Recipients, (Under Publication).