

UNIVERSITÀ DEGLI STUDI DELL'INSUBRIA

PHD PROGRAM IN METHODS AND MODELS FOR
ECONOMIC DECISIONS

UNIVERSITAT JAUME I

DOCTORAL PROGRAM IN ECONOMICS AND BUSINESS

**Essays in Behavioural and
Experimental Economics:
Emotions, Uncertainty and
Cooperation**

Author:

Matteo M. MARINI

Supervisors:

Prof. Ivan MOSCATI

Prof. Aurora GARCÍA-GALLEGO

February 26, 2019



UNIVERSITÀ DEGLI STUDI DELL'INSUBRIA

PHD PROGRAM IN METHODS AND MODELS FOR
ECONOMIC DECISIONS

UNIVERSITAT JAUME I

DOCTORAL PROGRAM IN ECONOMICS AND BUSINESS

Essays in Behavioural and Experimental Economics: Emotions, Uncertainty and Cooperation

Memoria presentada por Matteo M. Marini para optar al grado de doctor en régimen de
cotutela por la Università degli Studi dell'Insubria y por la Universitat Jaume I

Author:

Matteo M. MARINI

Supervisors:

Prof. Ivan MOSCATI

Prof. Aurora GARCÍA-GALLEGO

February 26, 2019



Contents

Acknowledgements	3
Preface	4
Prefacio	6
1 Incidental emotions and risk-taking: An experimental analysis	8
1.1 Introduction	8
1.2 Literature review	9
1.3 Experimental design	12
1.4 Results	16
1.4.1 Manipulation check	16
1.4.2 Overall analysis	18
1.4.3 Discussion	22
1.5 Conclusions	23
Appendix 1A. Experimental instructions	25
2 20 years of emotions and risky choices in the lab. A meta-analysis	27
2.1 Introduction	27
2.2 Related literature	29
2.3 Data and methodology	31
2.3.1 Effect size and identification of studies	32
2.3.2 Moderator variables	35
2.4 Analysis and results	39
2.5 Conclusions	46
Appendix 2A. Papers included in meta-analysis	48

Appendix 2B. Data	50
3 Communication in a threshold public goods game with ambiguity: Anomalies and regularities	53
3.1 Introduction	53
3.2 Literature review	56
3.3 Experimental design and hypotheses	59
3.3.1 The threshold public goods game	60
3.3.2 The Ring-test and the final questionnaires	63
3.3.3 Hypotheses	65
3.4 Results	67
3.4.1 Overall analysis	67
3.4.2 Group outcomes	72
3.4.3 Chat analysis and further insights	75
3.5 Conclusions	79
Appendix 3A. Experimental instructions	81

Acknowledgements

In connection with the first chapter, I acknowledge Università Politecnica delle Marche for financial support (project No. 7080 “Ricerca Scientifica di Ateneo”, year 2012, scientific coordinator: Alberto Russo) and the technical staff of the Faculty of Economics “Giorgio Fuà”, notably Daniele Ripanti, for flawless assistance. I am truly grateful to Annarita Colasante and Alberto Russo for having instilled confidence and interest in the experimental method in me, during a career stage where failure is a common occurrence. Moreover, I wish to thank Ivan Moscati, the academic staff and PhD students of Università degli Studi dell’Insubria for active participation in the workshop organized by the same university in Varese (March 27, 2017). Likewise, I am thankful to the audience of the seminar held at Universitat Jaume I in Castellón (May 23, 2017), as well as to those who attended the Young Economists’ Meeting in Brno (June 1-2, 2017) and the 9th SEET workshop in Lecce (February 23-25, 2018) for helpful remarks. Last but not least, my heartfelt thanks to Maria Bigoni and Nikolaos Georgantzis for insightful and constructive comments that produced substantial improvement in the analysis.

Regarding the genesis of the second chapter, I wish to thank both my supervisors who throughout the PhD program have encouraged me to come up with own topics and to deepen my knowledge of Experimental Economics and Econometrics, so that I could become autonomous in all the phases of research process.

As far as the third chapter is concerned, I gratefully acknowledge the Spanish *Ministerio de Economía y Competitividad* (grant ECO2015-68469-R) as well as the technical staff of the Laboratorio de Economía Experimental for careful assistance. Aurora García-Gallego and Luca Corazzini deserve a special mention for having dispensed valuable advice whenever the path to completion seemed to be particularly steep. In addition, I wish to thank the audience of the PhD workshops held at Università degli Studi dell’Insubria in Varese (May 3, 2018) and at Universitat Jaume I in Castellón (May 14, 2018) for active participation, and I am truly grateful to the participants in the 11th Maastricht Behavioral Experimental Economics Symposium in Maastricht (June 4, 2018), in the 15th Workshop on Social Economy for Young Economists in Forlí (June 8-9, 2018), in the 9th International Conference of the French Association of Experimental Economics in Nice (June 14-15, 2018) and in the Vechta Workshop on Social Behavior in Vechta (January 24-25, 2019) for helpful remarks.

Finally, let me thank Università degli Studi dell’Insubria and Universitat Jaume I for having stipulated the agreement for joint supervision, two anonymous reviewers for having brought considerable improvement to a previous version of this doctoral thesis, as well as three committee members for their admirable efforts at taking part in the final examination held in Varese (February 26, 2019).

Preface

In the present doctoral thesis I apply the experimental method in the context of two lines of research in Economics, the former investigating the role of incidental emotions in decision making under risk, whereas the latter shedding light on the impact of communication on cooperation. The dissertation consists of three sections.

In the first chapter I conduct a laboratory experiment in order to study the effect of incidental sadness and happiness on risky decision making. An emotion induction procedure is the treatment variable of a between-subjects design where two sessions aim at eliciting either sadness or happiness, respectively. Two further groups are characterized by neutral conditions and serve as baseline. After a manipulation check verifies the validity of the induction procedure, I use a multiple price list à la Holt and Laury (2002) to elicit individual risk preferences in the context of a lottery-choice task. The analysis reveals that both sadness and happiness promote greater risk aversion with respect to neutral conditions, a result which might be moderated by the risk elicitation task. Therefore, as effective explanation I propose the theory of ego depletion, whereby regulating emotions so as to subsequently process information consumes a limited self-control resource, which is needed to take risks as well.

The second chapter is a meta-analysis of experimental studies on the same topic, so as to explain traditional heterogeneity of outcomes in the field. After performing an advanced search in Google Scholar and filtering out studies that do not match a list of selection criteria, I include 16 studies from which 46 observations are drawn at the treatment level. At this point, I code a set of moderator variables representing experimental protocols and calculate Cohen (1988)'s d effect size as dependent variable of a weighted least squares (WLS) regression where larger studies are given more weight. Among the results, which are robust to different techniques for computing standard errors, I find that emotions induce higher risk aversion when a multiple price list à la Holt and Laury (2002) is used in place of stated preferences methods, as well as in case the risk elicitation task is framed as an investment decision instead of an abstract choice. Given the variety

of procedures employed in this type of experiments and in the absence of a tailor-made game to answer such research questions, I recommend faithful study replication as preferential path in order to investigate the influence of emotions on risky decision making and ensure comparability.

The third chapter offers evidence on the impact of communication on the provision of public goods whose quality is uncertain. I run a laboratory experiment with two treatments, where the control variable is pre-play communication in the form of unrestricted text chat. A binary threshold public goods game with four-person groups and threshold of three is at the core of the design, the main novelty lying in the provision mechanism with ambiguity. Moreover, a private signal for the actual value of the public good is provided, before the contribution decision. In accordance with related literature, I find that communication significantly increases public good provision by reducing inefficiency that comes from wasteful undercontribution. Nevertheless, the players in the chat treatment seem to neglect the free-rider issue and often end up overcontributing, in contrast with previous scientific findings. After chat analysis, I propose the pursuit of symmetric payoffs within the group as original explanation of the massive overcontribution, in addition to group identity generated by the partner matching and the common fate hypothesis. Since the players prefer to minimize ambiguity than to maximize the group earnings, I finally speculate that under uncertainty satisficing is more salient than optimizing.

Prefacio

En esta tesis doctoral se aplica la metodología experimental en el contexto de dos líneas de investigación en Economía. La primera línea investiga el papel de las emociones casuales en la toma de decisiones en situaciones arriesgadas. La segunda línea de investigación trata de aportar luz al impacto de la comunicación sobre la cooperación. La tesis consta de tres capítulos.

En el primer capítulo se realiza un experimento de laboratorio con el fin de estudiar el efecto de las emociones de tristeza y felicidad, sentidas como casuales no relacionadas con la decisión propiamente dicha, sobre la toma de decisiones bajo riesgo. El proceso de inducción de una emoción es la variable tratamiento de un diseño entre-sujetos en el que hay dos sesiones cuyos objetivos son, respectivamente, elicitación de tristeza y felicidad. Dos grupos adicionales se caracterizan por condiciones neutrales y sirven de tratamiento base. Tras cierta manipulación que permite verificar la validez del proceso de inducción, se utiliza una lista múltiple de precios a la Holt and Laury (2002) para elicitación de las preferencias individuales frente al riesgo en el contexto de una tarea de elección de loterías. El análisis revela que los dos tipos de emociones, tanto tristeza como felicidad, suponen mayor aversión al riesgo en relación a condiciones neutrales, un resultado que puede moderarse por la tarea de elicitación de actitud frente al riesgo. Por lo tanto, como explicación persuasiva se propone la teoría del agotamiento del ego, que establece que regular las emociones a fin de procesar información consume un recurso limitado, el autocontrol, también necesario para tomar decisiones más arriesgadas.

El segundo capítulo es un metanálisis de estudios experimentales sobre el tema analizado en el primer capítulo, con el objetivo de explicar la heterogeneidad tradicional de los resultados en la literatura sobre el tema. Tras realizar una profunda búsqueda en Google Scholar y tras dejar fuera estudios que no casan con una lista de criterios de selección, se incluyen 16 estudios de los que se extraen 46 observaciones a nivel de tratamiento. Tras ello, se codifican un conjunto de variables moderadoras que representan protocolos experimentales y se calculan los efectos tamaño d de Cohen (1988) como variable dependiente de un modelo de regresión por mínimos cuadrados ponderados (WLS) en el que se ha dado mayor peso a estudios con más obser-

vaciones. Un resultado que es robusto ante diferentes métodos de cálculo del error estándar es que las emociones inducen mayor nivel de aversión al riesgo cuando la elicitación se realiza a partir de una lista múltiple de precios a la Holt and Laury (2002) en vez de con métodos de preferencias declaradas. También ocurre cuando la tarea de elicitación de la actitud frente al riesgo se etiqueta como una decisión de inversión en vez de una elección abstracta. Dada la variedad de procedimientos utilizados en este tipo de experimentos y en la ausencia de un juego hecho a medida para contestar a las cuestiones planteadas, se recomienda, en aras de la comparabilidad, la réplica fiel como pauta preferencial con el objetivo de investigar la influencia de las emociones en la toma de decisiones bajo condiciones arriesgadas.

El tercer capítulo ofrece evidencia sobre el impacto de la comunicación en la provisión de bienes públicos cuya calidad es incierta. En este trabajo se diseña un experimento de laboratorio con dos tratamientos, siendo la variable de control la comunicación previa a la toma de decisiones a través de un chat online de texto libre. El núcleo del diseño es un juego de bienes públicos con decisión binaria, grupos de cuatro jugadores y umbral de tres. La principal novedad del diseño está en el mecanismo de provisión con ambigüedad. Además, la provisión del bien público depende de una señal privada que se envía antes de que se tome la decisión de contribuir. Acorde con la literatura relacionada con el tema, se encuentra que la comunicación incrementa de forma significativa la provisión del bien público reduciendo la ineficiencia que proviene de una contribución derrochadora e ineficiente. En contraste con investigaciones anteriores, sin embargo, los sujetos en el tratamiento con chat parecen ignorar el asunto del polizón y, muy a menudo, acaban sobre-contribuyendo. Tras un análisis exhaustivo del chat, se propone la búsqueda de pagos simétricos dentro del grupo como explicación a una masiva sobre-contribución, además de una identificación con el grupo generada por el emparejamiento fijo y la hipótesis del destino común. Dado que los jugadores prefieren minimizar la ambigüedad a maximizar las ganancias del grupo, se especula en definitiva que, en condiciones de incertidumbre, se prioriza satisfacer a optimizar.

Chapter 1

Incidental emotions and risk-taking: An experimental analysis

1.1 Introduction

At the present time, choice behavior is commonly regarded as the product of two interacting processes, say, the former more deliberative and logical, whereas the latter more impulsive and driven by emotions (Loewenstein et al., 2015). Despite this, scholars in the social sciences have been firstly interested in the development of the rational part of the story, this occurring notably in the field of risky decision making, if one considers the long-standing expected utility theory (EUT). The trend has been inverted only in the second half of the past century, when the outcomes of early behavioral experiments revealed the failure of mainstream theories and renewed intellectual curiosity about the affective sphere. Since then, researchers have lavished increasing endeavors to discover what sort of influence emotions exert on risk attitude, by virtue of far-reaching implications which range from organizational contexts (Andrade and Ariely, 2009) to individual investment decisions (Cohn et al., 2015).

The present chapter intends to make an empirical contribution to a recent literature whose outcomes are still inconclusive and inconsistent with each

other. In order to investigate the effect of incidental sadness and happiness on risky choices, I conduct a laboratory experiment with an emotion induction procedure as treatment variable and a between-subjects design. As to its structure, two treatments aim at inducing either sadness or happiness, respectively, whereas two further groups are characterized by neutral conditions and serve as baseline. After a manipulation check verifies the validity of the induction procedure, I use a multiple price list (MPL) à la Holt and Laury (2002) to elicit individual risk preferences in the context of a lottery-choice task.

The analysis reveals that both sadness and happiness promote greater risk aversion than neutral conditions, a result which is in line with Drichoutis and Nayga Jr (2013) but not supported by the Affect Infusion Model and the Mood Maintenance Hypothesis, the two benchmark theories in the area of interest. The circumstance that the above-mentioned authors in turn employed a MPL highlights the concern about the risk elicitation method as potential moderator for the emotional influence on risky decision making, this being worth further investigation.

To conclude, as effective explanation I suggest that the experimental subjects, in an attempt to regulate sadness (or happiness) and carry out the risk elicitation task, end up being ego-depleted. Indeed, there is evidence that people under ego depletion become more pessimistic and, consequently, might regard their own chances as less favorable (Fischer et al., 2007), thereby avoiding risky choices.

The remainder of the chapter is organized as follows. After Section 1.2 reviews the related literature, the experimental design is outlined in Section 1.3. The results are introduced in Section 3.4, which includes the manipulation check as well as the overall analysis, before I discuss consistent explanations. Section 3.5 concludes.

1.2 Literature review

The academic interest in the role of emotions in risky decision making dates from the last decades of the past century, with the topic catching mainly the attention of psychologists. As fruits of these early efforts, two contra-

dictory theories still prevail in the literature, that is, the Affect Infusion Model (AIM) (Forgas, 1995) and the Mood Maintenance Hypothesis (MMH) (Isen and Patrick, 1983). On the one hand, the former suggests that positive (negative) affect leads to perceive a given situation as more (less) favorable, thereby fostering relatively risk-seeking (prudent) behavior. On the other hand, the latter posits that, when feeling good (bad), people are protective of (impatient to change) their own positive (negative) affective states and consequently take less (more) risks.

After the pioneering article by Elster (1998) spotlights the debate in Economics as well, traditional observational studies using proxies for emotions (Kliger and Levy, 2003; Kamstra et al., 2003) gradually make way for papers based on experiments, which definitely ensure greater reliability when it comes to identifying the causality nexus between affect and ensuing risky choices. At this point, results from early experiments point out that anxiety has different impact on risk attitude in comparison with sadness (Raghunathan and Pham, 1999), as well as fear does not produce the same effect as anger (Lerner and Keltner, 2001). This implies that valence, which refers to the degree that an emotion is positive or negative, represents a relevant factor to predict risk-seeking behavior but cannot be the only one, as Zeelenberg et al. (2008) notice. Therefore, the same authors invite researchers to focus on specific emotions when it comes to addressing similar research questions.

At the same time, the increasing use of procedures artificially inducing affective states in the context of experiments involves greater production of papers dealing with emotions, rather than moods. Indeed, in both Economics and Psychology the consensus is that *emotions* are short feelings triggered by a specific stimulus, whereas *moods* persist longer but their origin is undefined, with the term *affect* acting as umbrella concept and encompassing both the previous definitions (Robbins and Judge, 2012; Lerner et al., 2015). Moreover, given that the emotions elicited during such experiments are often unrelated to the following choice to be made, *incidental* emotions are more frequently investigated than *integral* emotions, the case where, on the contrary, affect is linked to the outcome of the decision at hand (Rick and Loewenstein, 2008).

To date, the AIM and the MMH have continued to serve as benchmarks for ongoing research, even though these psychological theories still rely on a valence-based approach. Manuscripts uniformly endorse either the AIM (Grable and Roszkowski, 2008; Stanton et al., 2014) or the MMH (Kliger and Levy, 2003; Kim and Kanfer, 2009), irrespective of the methodology used by the authors. Other studies come to intermediate conclusions, in the sense that do not fully support either theory (Drichoutis and Nayga Jr, 2013; Conte et al., 2018).

As Treffers et al. (2016) propose, the mixed evidence might be due to the fact that according to the experimental protocols in the field of Psychology subjects do not need to be remunerated, differently from the practices in Economics. Since risk tolerance is a socially desirable quality, participants may want to appear risk-lover when stakes are hypothetical and misreporting risk attitude is costless (Camerer and Hogarth, 1999). Likewise, given the common mild evidence for treatment effects, publication bias could also play a role in case the magnitude of effect sizes is so small that it prevents papers from being published.

In particular, the inconsistencies might be ascribed to the heterogeneity of risk elicitation methods, considering the variety of procedures adopted in the literature. Indeed, this array ranges from simple self-reported measures of risk propensity (Yuen and Lee, 2003; Chou et al., 2007), usually obtained through non-incentivized questionnaires, to more sophisticated tasks estimating risk parameters of models with assumptions on the utility function (Drichoutis and Nayga Jr, 2013; Fehr-Duda et al., 2011). In an attempt to guide the reader to an informed choice, Charness et al. (2013) illustrate advantages and disadvantages of the most commonly used methods, reaching the conclusion that the decision depends on the research question and the characteristics of the sample population. This insight is shared by Crosetto and Filippin (2016), who in a theoretical and experimental assessment of four risk elicitation methods explain the between-tasks inconsistencies partly by construction and in part because of behavioral artifacts. Among the latter, for instance, I could mention the context of risky decision making, since it is well-established that familiarity with the domain of choices reduces risk perception (Hanoch et al., 2006).

Acknowledging that solving the issue of study-to-study variability in previous findings goes beyond the scope of this chapter, now I aim to make an empirical contribution to the discussion and outline the experimental design.

1.3 Experimental design

The experiment is programmed by means of the software z-Tree (Fischbacher, 2007) and held at the Laboratory of the Faculty of Economics of Università Politecnica delle Marche (Ancona, Italy), involving 123 students (59 females) recruited from the university campus thanks to the system ORSEE (Greiner, 2015).

As summarized in Table 3.1 and Table 3.2, the subjects are randomly assigned to four treatments (here coinciding with sessions) that differ from one another in the control variable, namely, the emotion induction procedure. In more detail, the design consists of two treatments which aim at eliciting either sadness or happiness, respectively, in addition to two treatments serving as baseline and obtaining neutral conditions with two different methodologies.

Table 1.1: Treatments

	S-treatment	H-treatment	Baseline	
Subjects	35	30	30	28
Targeted emotion	Sadness	Happiness	Neutrality	Neutrality
Induction proced.	Sad stim.	Happy stim.	Neutral stim.	-

Except for the induction procedure, each session follows the same scheme in such a way that any difference in risk attitude captured by the risk elicitation task can be ascribed to the targeted emotion, which is triggered by the treatment variable. In the end, one of the subject's choices in the task is randomly selected and paid out for real, this payment method being equivalent to remunerating all the decisions (Cubitt et al., 1998; Hey and Lee,

Table 1.2: Structure of each session

Order	Phase
1	Socio-demographic questionnaire
2	Induction procedure (<i>treatment variable</i>)
3	Manipulation check
4	Risk elicitation task
5	Payment feedback

2005). Each session lasts approximately 40 minutes and the mean earnings amount to 8 euro, including a show-up fee of 3 euro.

At first, I let the participants take their seats in the cubicles so that the instructions are immediately explained aloud, with a focus on the monetary incentives provided. Hence, the subjects are supposed to fill in a socio-demographic questionnaire which, in addition to a number of filler items, controls for individual characteristics potentially interacting with risky decision making (gender, age, and willingness to take risks).¹

At this point, I invite the participants to put on the headphones, thereby conveying an audiovisual stimulus which is treatment-specific. Indeed, drawing inspiration from a battery of techniques validated by Westermann et al. (1996), a combination of music and pictures is used in order to elicit the targeted emotions with greater efficacy than either method taken separately (Mayer et al., 1995):

- in the *S-treatment* (sad treatment) I broadcast a couple of minutes from the musical piece “Polymorphia” by Krzysztof Penderecki along with a slideshow of images from concentration camps;
- in the *H-treatment* (happy treatment) the selected track is Leonard Bernstein’s “Mambo”, whose video is shown in the passionate perform-

¹The rationale behind the placement of the questionnaire at the outset lies in the question about willingness to take risks, which represents a dispositional measure of risk-seeking behavior and, in order to be meaningful, needs to be asked before the induction procedure.

Table 1.3: The multiple price list in the version by Holt and Laury (2002)

No.	Option A		Option B	
1	1/10 of 2.00,	9/10 of 1.60	1/10 of 3.85,	9/10 of 0.10
2	2/10 of 2.00,	8/10 of 1.60	2/10 of 3.85,	8/10 of 0.10
3	3/10 of 2.00,	7/10 of 1.60	3/10 of 3.85,	7/10 of 0.10
4	4/10 of 2.00,	6/10 of 1.60	4/10 of 3.85,	6/10 of 0.10
5	5/10 of 2.00,	5/10 of 1.60	5/10 of 3.85,	5/10 of 0.10
6	6/10 of 2.00,	4/10 of 1.60	6/10 of 3.85,	4/10 of 0.10
7	7/10 of 2.00,	3/10 of 1.60	7/10 of 3.85,	3/10 of 0.10
8	8/10 of 2.00,	2/10 of 1.60	8/10 of 3.85,	2/10 of 0.10
9	9/10 of 2.00,	1/10 of 1.60	9/10 of 3.85,	1/10 of 0.10
10	10/10 of 2.00,	0/10 of 1.60	10/10 of 3.85,	0/10 of 0.10

ance by Gustavo Dudamel and the Simón Bolívar Youth Orchestra;

- in the related two sessions, the *baseline* seeks to induce neutral conditions respectively by means of (i) a combination of pictures of animals and an excerpt from the first movement of the very well-known “Symphony No. 40” by Wolfgang Amadeus Mozart, piece already validated by Västfjäll (2002); (ii) the absence of any induction procedure.

In accordance with the protocols in the field, at this stage the inducement of the targeted emotions is proved through a manipulation check which consists in asking participants to report their own affective state, this being presented in the results section.

Afterwards, the subjects carry out a risk elicitation task in the form of multiple price list (MPL), a format adopted for the first time by Binswanger (1980) and popularized by Holt and Laury (2002) in the version reported in Table 1.3.²

²Apart from that one illustrated in the text, the design includes two successive multiple price lists which are not part of the research of this work, and accordingly will not be further considered.

In a nutshell, the original task is based on a ten-row list where the subjects are supposed to make ten decisions in succession between paired lotteries, with one choice being played out for real as salient incentive. In each row, Option A is less variable and thus considered safer than Option B, given that the outcome probabilities are the same. Moreover, since the probabilities to win the highest amounts increase with the choice number, at the same time Option B becomes more and more inviting so that by the last row of the table even the most risk-averse person is expected to switch from the safer Option A to Option B. Speculating that people behave in line with expected utility theory (EUT), each choice corresponds to an interval of the parameter measuring risk aversion under the assumption of a specific functional form, that is, the Constant Relative Risk Aversion (CRRA) utility function. Therefore, the switch point captures the individual degree of risk aversion. Otherwise, as the authors notice, greater generality can be obtained by considering the amount of safe choices, which in Table 1.3 corresponds to the number of Options A.

In an attempt to correct some drawbacks of the task according to the suggestions in Abdellaoui et al. (2011), my design modifies the version by Holt and Laury (2002) as follows.

Instead of comparing lotteries, the subjects are supposed to choose between lotteries and sure payoffs, this leading to superior tractability. As to the lotteries, one of the two outcomes is set to 0 for the sake of clarity (Shupp et al., 2013). Finally, I aim at increasing generality and replace the probability scale with an outcome scale, so as to take into consideration also non-expected utility behaviors where people evaluate probabilities non-linearly.

The resulting MPL is contained in Table 1.4, where an individual switching from B to A between choices 6 and 7 exhibits risk neutrality. If one switches earlier (later), instead, she reveals herself to be risk-averse (risk-lover).

As said, in the end one of the ten choices is randomly selected and played out, so that the earnings are converted into euro for the final payment.

Given the inconsistencies in previous findings related to sadness and happiness, I have no prior beliefs about the direction of the effects on risk attitude.

Table 1.4: The multiple price list in my experimental design

No.	Option A	Option B	Exp. payoff difference
1	100	1/2 of 300, 1/2 of 0	- 50
2	110	1/2 of 300, 1/2 of 0	-40
3	120	1/2 of 300, 1/2 of 0	-30
4	130	1/2 of 300, 1/2 of 0	-20
5	140	1/2 of 300, 1/2 of 0	-10
6	150	1/2 of 300, 1/2 of 0	0
7	160	1/2 of 300, 1/2 of 0	10
8	170	1/2 of 300, 1/2 of 0	20
9	180	1/2 of 300, 1/2 of 0	30
10	190	1/2 of 300, 1/2 of 0	40

1.4 Results

Now I present the results of the experiment within a three-parted section. The structure of the manipulation check is introduced at the beginning, before I perform an overall analysis and finally discuss consistent explanations in a devoted subsection.

1.4.1 Manipulation check

Before moving to the analysis, I need to go through a couple of preparatory steps. The manipulation check allows me to verify that the targeted emotions were correctly induced, this implying that (i) the two alternative methodologies to induce neutrality actually led to the same outcome, and (ii) both sadness and happiness were properly elicited in the respective treatments.

Hence, between the emotion induction procedure and the risk elicitation task I asked subjects to report a self-evaluation of the current affective state by using 6-point Likert scales for a number of filler adjectives, in addition to those ones of interest (Matell and Jacoby, 1971).

Table 1.5 shows that the adjectives were treatment-specific, thereby con-

Table 1.5: Emotion-related questions after the induction procedure

	S-treatment	H-treatment	Baseline	
How much are you from 0 to 5?	Sad	Happy	Sad	Sad
	Worried	Carefree	Happy	Happy
	Scared	Serene	-	-
	Doubtful	Determined	-	-

tributing to the emotion inducement. At the same time, potential demand effects are not a concern, given that the direction of the emotional impact on risk attitude is not intuitive. Furthermore, since the subjects were free to respond “0” and signal indifference, the procedure permitted to collect reliable answers (Campos-Vazquez and Cuilty, 2014), as well as to make pairwise comparisons with the baseline.

For the sake of interpretation, the 6-point Likert scale is dichotomized so that each individual is classified as under the influence of emotion just in case her self-evaluation ranges from 3 to 5 with respect to the corresponding adjective. Accordingly, a score between 0 and 2 entails absence of the affect in question.

Table 1.6: Emotions in the two baseline sessions

	With stimulus	Without stimulus
Sad	2	6
Non-sad	28	22
<i>Total</i>	30	28
Happy	22	20
Non-happy	8	8
<i>Total</i>	30	28

Absolute frequencies.

As displayed in Table 1.6, by so doing the 7% of subjects (2 out of 30) in

the baseline session with neutral stimulus are categorized as *sad*, with this percentage not differing from the 21% (6 out of 28) in the group without induction procedure according to a two-tailed Fisher’s exact test ($p = 0.138$). Likewise, the distribution of *happy* participants in the two sessions is statistically similar (73% vs 71%), as the same test confirms ($p = 1.000$). This implies that the neutral induction procedure is validated and allows me to handle the two baseline groups as a whole in the remainder of the analysis.

Table 1.7: Comparing treatments and baseline

	S-treatment	H-Treatment	Baseline
Sad	16	-	8
Non-sad	19	-	50
<i>Total</i>	35	-	58
Happy	-	27	42
Non-happy	-	3	16
<i>Total</i>	-	30	58

Absolute frequencies.

Moreover, Table 1.7 reports that on the one hand, the subjects in the S-treatment are classified as *sad* in 46% of cases (16 out of 35), which according to a one-tailed Fisher’s exact test test is actually higher than the 14% (8 out of 58) in the baseline, with significance at the 1% level. On the other hand, the proportion of *happy* participants in the H-treatment is significantly greater than in the neutral sessions (90% vs 72%), this resulting from the same test ($p < 0.05$).

Therefore, I conclude that the targeted emotions were correctly induced, and in the next subsection I am able to attribute possible differences in risk attitude to the dissimilar affect experienced at the time of decision making.

1.4.2 Overall analysis

Before presenting the results, in particular I highlight the coding of two variables:

- *willingness to take risks (WTTR)*, the answer to the following question contained in the initial questionnaire: “You are offered the opportunity to acquire a security which, with the same probability, allows you either to gain 1000 euro or to lose the invested sum. How much are you willing to pay for such a security?” (Guiso and Paiella, 2008).
- *number of risky choices*, the individual amount of risky decisions in my version of the multiple price list (Table 1.4), where each Option A is considered *safe* and each Option B is classified as *risky*. This number ranges from 0 to 10, with higher values corresponding to greater risk tolerance.

Table 1.8: Summary statistics by treatment

	S-treatment	H-Treatment	Baseline
No. of risky choices	4.66	5.13	5.79
Female	0.43	0.43	0.53
Age	22	22	23
WTTR	80	80	100
<i>No. of subjects</i>	35	30	58

Means for number of risky choices; relative frequencies for gender; medians for age and willingness to take risks (WTTR).

The related summary statistics are reported in Table 1.8 in addition to some demographics. According to a battery of Kruskal-Wallis tests, the three groups are not drawn from the same population with respect to the median, when I look at individual characteristics (p -values equal to 0.519, 0.299, and 0.528 for female, age and willingness to take risks, respectively). This confirms that the subjects have been randomly allocated across treatments. Instead, the test finds significant differences as to the amount of risky choices ($p < 0.01$), this preliminary result encouraging towards further analyses.

Hence, by considering the number of risky choices as outcome variable of an ordered logit model, I follow the original approach by Holt and Laury (2002) and avoid dropping the 49 subjects out of 123 (40%) who switched

Table 1.9: Explaining risky choices

Dependent variable: No. of risky choices		
	(1)	(2)
<i>Sadness</i>	-1.257*** (0.409)	-1.279*** (0.398)
<i>Happiness</i>	-0.773* (0.415)	-0.772* (0.399)
Female	-0.050 (0.324)	
Age	-0.064 (0.078)	
WTTR	0.001 (0.001)	
LR χ^2	13.01	11.13
Prob > χ^2	0.023	0.004
Pseudo R-squared	0.027	0.023
No. of subjects	123	123

Coefficient estimates from the ordered logit model, with standard errors in parentheses. *Sadness* and *happiness* are the treatment dummies.

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

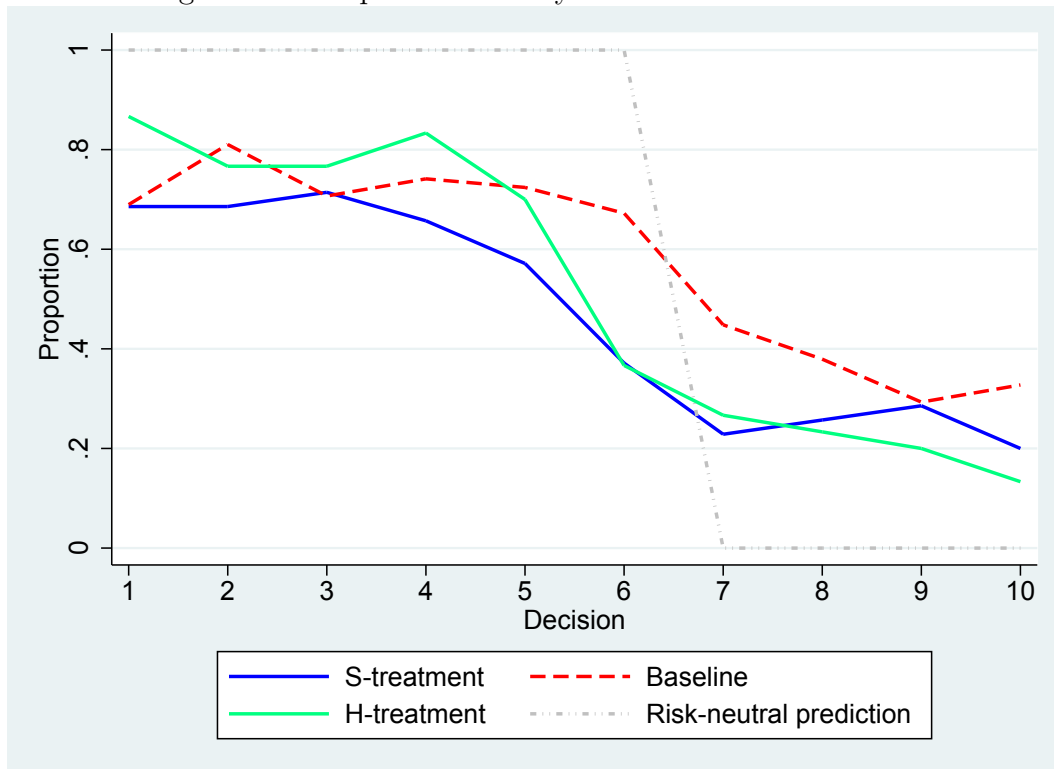
back and forth in the MPL.³

As shown in Table 1.9, the coefficients of the treatment dummies *sadness* and *happiness* are negative and statistically significant, this meaning that both emotions fostered risk aversion with reference to the baseline. Regarding this outcome, which can be appreciated in Figure 1.1, I observe a strong effect of sadness ($p < 0.01$) but only mild evidence as to happiness ($p = 0.06$). A

³All the data, tables and figures reported in this chapter use the full sample of observations. Nevertheless, concerning the decision to include such inconsistent subjects in the analysis, there is no agreement in the literature. For instance, Crosetto and Filippin (2016) regard this choice as not methodologically sound. In order to tackle this concern, I repeat the analysis by restricting the attention to those subjects who never switched from Option A to Option B, but the medians by treatment are unchanged. An Epps-Singleton test fails to reject the hypothesis of no differences in the distribution of the number of risky choices between the restricted sample and the full one, this holding true for all three treatments.

reasonable explanation lies in the fact that the inducement of happiness was slightly less effective than the sadness elicitation ($p < 0.05$ vs $p < 0.01$), thus leading to smaller effect size when it comes to the number of risky choices. Accordingly, the limited effect magnitude associated with happiness would require greater sample size in order to be detected, as power analysis confirms.⁴

Figure 1.1: Proportion of risky choices in each decision



Finally, the findings are not driven by any of the individual characteristics, so that in column (2) I maintain only the treatment dummies. This second specification emphasizes that the effects are genuine.

⁴When I compare the H-treatment with the baseline, power is equal just to 0.36, a value which appears to be distant from the benchmark of 0.80 (Moffatt, 2015). On the contrary, the comparison between S-treatment and baseline reaches the satisfactory level of 0.78.

1.4.3 Discussion

Searching for explanations consistent with the results, the long-standing theories of the Affect Infusion Model (AIM) (Forgas, 1995) and the Mood Maintenance Hypothesis (MMH) (Isen and Patrick, 1983) do not fully account for the fact that both the S-treatment and the H-treatment promoted risk aversion with respect to neutral conditions.

On the one hand, the conjectures of the AIM are in line with (i) the relatively cautious choices made under sadness, yet fail to explain (ii) such a behavior under happiness. On the other hand, the MMH supports (ii) but is not consistent with (i).

Nevertheless, my findings are not actually unprecedented, given that Drichoutis and Nayga Jr (2013) come to the same conclusions by employing a multiple price list (MPL) à la Holt and Laury (2002) as well. The circumstance renews the idea of the risk elicitation method as possible moderator for the influence of emotions on risky decision making, this being worth further investigation.

As effective explanation, the results of the chapter are coherent with the argument that the psychological phenomenon known as *ego depletion* (Baumeister et al., 1998) played a role during risky decision making. According to this theory, suppressing emotions, processing information and making active choices are only some of a number of mental activities consuming a limited self-regulatory resource, which can be seen as a kind of energy. Once this resource is even partly used, subsequent performance on self-control tasks will be impaired.

Regarding the consequences of ego depletion on risky decision making, Unger and Stahlberg (2011) have recently organized the area by distinguishing two types of risk behavior. If the decision task takes the shape of a gamble without perceived control over the final outcomes, such as playing the lottery or roulette, then ego-depletion conditions are expected to increase risk-taking. Whereas, greater risk aversion is predicted when it comes to a decision scenario where people are more responsible. For instance, as the authors themselves suggest, this is the case of “economic decisions or the selection of various options that differ from each other with regard to out-

come probabilities and pay-offs”, so long as both the negative and positive consequences of the decision materialize immediately. Otherwise, reduced self-control fosters risk-seeking behavior, such as in the successive choice to smoke, whose emotional rewards are instantaneous but negative effects arise only in the distant future.

Basically, the conclusions from this psychological strand are compatible with the scenario whereby my experimental subjects, in an attempt to regulate sadness (or happiness) and carry out the risk elicitation task, ended up being ego-depleted. Given that the MPL à la Holt and Laury (2002) entailed reasoning about outcome probabilities and pay-offs, the resulting risk aversion is in line with previous literature.

Such a behavior under ego depletion is further justified by strategic thinking, if one takes her own reduced self-control into account and thus considers risk-taking as an emotional threat. Finally, there is also evidence that people in this condition exhibit “a less-optimistic sense of their own abilities, a lower sense of subjective control, and less-optimistic expectations about their future” (Fischer et al., 2007).

1.5 Conclusions

In this chapter I conducted a laboratory experiment in order to investigate the effect of incidental emotions on risky decision making. On the one hand, the work found a source of motivation in the large number of fields which could benefit from the results, from organizational contexts to individual investment decisions. On the other hand, I aimed to make an empirical contribution to a recent literature whose outcomes are still inconclusive and inconsistent with each other.

The analysis revealed that both sadness and happiness promoted greater risk aversion with respect to neutral conditions, a result which is in line with Drichoutis and Nayga Jr (2013) but not supported by the AIM and the MMH, the two benchmark theories in the area of interest. The circumstance that the above-mentioned authors in turn used a multiple price list à la Holt and Laury (2002) reaffirmed the concern about the risk elicitation method as potential moderator for the emotional impact on risky decision making,

this being worth further investigation.

As effective explanation, I suggested that the experimental subjects, in an attempt to regulate sadness (or happiness) and carry out the risk elicitation task, ended up being ego-depleted. Indeed, there is evidence that people under ego depletion become more pessimistic and, consequently, might regard their own chances as less favorable (Fischer et al., 2007), thereby avoiding risky choices.

To conclude, given the enduring inconsistencies in findings, I recognize the issue of study-to study variability as major challenge for future academic efforts.

Appendix 1A. Experimental instructions

[The following instructions are translated from Italian. They were displayed on monitors and were the same across treatments, except for screens 4a, 4b, and 5.]

SCREENS 1 and 2. You are going to take part in a game aiming to study individual behavior. During the game you will be asked to fill in a questionnaire and then to make decisions. Throughout the session, the amounts displayed are expressed in the form of experimental currency units (ECU), your final profit being converted into euro in the end. Moreover, you will be able only to observe your own choices, without information about the other participants' decisions. Anyway, you are not expected to interact with the others and, accordingly, your final earnings will be affected just by your own choices. We finally want to remind you that during the experiment it is forbidden to speak with the other participants as well as to use mobile phones. In case you need further clarification, please raise your hand and a member of the staff will come to your assistance.

SCREENS 3 and 4. *[Here participants could fill in the socio-demographic questionnaire containing the three variables used in the analysis in addition to four filler questions (FQ)]* Please answer the following questions:

1. How old are you?
2. What is your gender?
3. What is your father's profession? (FQ)
4. What is your mother's profession? (FQ)
5. What is your mother's level of education? (FQ)
6. You are offered the opportunity to acquire a security which, with the same probability, allows you either to gain 1000 euro or to lose the invested sum. How much are you willing to pay for such a security? *[Here participants could select any integer between 0 and 1000 euro]*
7. In general, how much do you think you are healthy? Use the following 0-10 rating scale. (FQ)

SCREENS 4a and 4b. *[These screens were present only in S-treatment, H-treatment, and baseline with neutral stimulus. Here participants were verbally invited to put on the headphones so that we could convey a treatment-specific stimulus according to the guidelines in Section 1.3]*

SCREEN 5. Describe your current affective state by using a 0-5 rating scale for the following adjectives, where "1" corresponds to "Not at all" and "6" to "Very much". *[The adjectives displayed in the S-treatment were "sad", "worried", "scared", and "doubtful". The adjectives displayed in the H-treatment were "happy", "carefree", "serene",*

and “determined”. The adjectives displayed in the two baseline treatments were “sad” and “happy”.]

SCREENS 6 and 7. In this game you are supposed to make ten decisions in succession between a sure amount and a lottery which allows you to win either 300 ECU or nothing with the same probability. Your actual payoff will be determined through the random draw of one of the ten decisions. In case a decision where you chose the lottery is selected, then the lottery is played out for real too. Such a payoff will be added to an initial endowment of 100 ECU.

SCREEN 8. [*Here participants could make the ten binary decisions by carrying out the MPL in the version illustrated in Table 1.4, whose last column was not shown.*]

SCREENS 9-17. [*Here participants carried out other two MPL which are not part of the research agenda of this work, and accordingly are not considered.*]

SCREEN 18. [*Here we showed the individual final profit to participants.*]

Chapter 2

20 years of emotions and risky choices in the lab. A meta-analysis

2.1 Introduction

Starting from early criticism of the paradigm of *homo oeconomicus*, understanding the role of emotions in the broad area of decision making has been one of the greatest academic goals, say, a multidisciplinary challenge which many researchers have been willing to take on especially in the last two decades, if one considers that yearly scholarly papers on the topic amounted to over 400 in 2013, but “doubled from 2004 to 2007 and again from 2007 to 2011” (Lerner et al., 2015).

Despite in its infancy, the field is already so vast that a number of different approaches can be chosen in order to join the debate. Fortunately, two recent articles facilitate the task and provide a complementary view of the literature (George and Dane, 2016; Lane, 2017), thereby also highlighting the strands where the results are particularly unclear.

The former manuscript goes through the papers appeared in *Organizational Behavior and Human Decision Processes* since the 1990s and distinguishes topical areas according to whether the emotions arise from the decision to be made in the future (*integral* emotions) or, instead, from an un-

related stimulus (*incidental* emotions). However, both these lines of research investigate the influence of affects on subsequent choices, with a further section being devoted to affective states as consequences of decision-making.

The latter review, in addition, analyzes interactive behavior separately from individual behavior, yet concentrates its efforts on happiness only. In this composite framework, as Lane (2017) himself notices, research on the impact of happiness on risky choices is definitely a field rife with unanswered questions and contradictory results, but “the literature may be reaching the stage where a meta-analysis would be a useful endeavour”. Therefore, the present work accepts this invitation and aims to explain traditional heterogeneity of outcomes in the area of interest by means of meta-analytic evidence.

After performing an advanced search in Google Scholar and filtering out studies that do not match a list of selection criteria, I include 16 manuscripts from which 46 observations are drawn at the treatment level. At this point, I code a set of moderator variables representing experimental protocols and calculate Cohen (1988)’s d effect size as dependent variable of a weighted least squares (WLS) regression where larger studies are given more weight.

In the analysis I find that emotions induce higher risk aversion when a multiple price list à la Holt and Laury (2002) is used in place of stated preferences methods, as well as in case the risk elicitation task is framed as an investment decision instead of an abstract choice. These results are robust to different techniques for computing standard errors and in line with related literature not inspecting affects. The regressions also show that the role of geographic differences is unclear, and stakes do not systematically influence behavior yet salient payments increase variability of outcomes, thereby contradicting solid evidence (Hertwig and Ortmann, 2001). Overall, anxiety turns out to be the only emotion leading to effects of the same sign, in the direction of greater risk aversion. At the same time, the data confirm that the valence-based approach should not be followed, since emotions of the same valence trigger significantly different risky choices.

To conclude, given the variety of procedures employed in this type of experiments and in the absence of a tailor-made game to answer such research questions, I recommend faithful study replication as preferential path in order

to investigate the influence of emotions on risky decision making and ensure comparability.

The remainder of the chapter unfolds as follows. Section 2.2 reviews the literature related to the meta-regression. Section 2.3 offers an extensive view of data and methodology before the analysis is carried out in Section 2.4, where I present the results as well. Section 2.5 concludes.

2.2 Related literature

Notably in use in the medical sciences, meta-analysis is a quantitative literature review which combines and evaluates empirical results through the methodological rigor of statistical techniques. In a nutshell, meta-regression analyses share the same limitations as narrative literature reviews but hold a relative advantage, namely, the criteria by which the studies are included and any misuse can be discussed on objective technical grounds (Stanley, 2001). Generally, for each selected paper a summary statistic is computed in the form of standardized mean difference or regression coefficient, this index capturing the effect magnitude and acting as dependent variable. Instead, the regressors represent characteristics of the original design which are deemed to be responsible for the study-to-study variation of the research findings. Hence, it is not surprising that this statistical tool rapidly proliferated in the branch of Experimental Economics (Zelmer, 2003; Oosterbeek et al., 2004; Engel, 2011; Johnson and Mislin, 2011; Capraro et al., 2018), where the features of designs can be easily coded.

In line with this tendency, the current chapter focuses on the experimental literature for a couple of reasons. As Lane (2017) suggests, laboratory experiments outperform other research methods on the opportunity to observe actual risk-seeking behavior rather than just self-reported attitude. Moreover, the controlled environment typical of the lab provides clear advantage over survey-based studies in identifying causality. Finally, the presence of an affect induction procedure allows to distinguish emotions from moods, which by nature do not arise from a triggering stimulus and hold different traits (Robbins and Judge, 2012; Lerner et al., 2015).

The strand of literature on how risk preferences are shaped by incidental

emotions is characterized by two opposing psychological theories, that is, the Affect Infusion Model (AIM) (Forgas, 1995) and the Mood Maintenance Hypothesis (MMH) (Isen and Patrick, 1983). The former predicts positive emotions to induce risky, and negative emotions safe choices, while the latter gives the inverse prediction. As previously mentioned, the topic has produced a number of papers with unclear results which do not fully support either theory, to such an extent that a meta-regression might be a worthwhile effort. The following studies, included in the analysis according to the selection criteria enumerated in Section 2.3, offer an accurate representation of such heterogeneity of outcomes.¹

Raghunathan and Pham (1999) find that sad subjects make relatively risky choices but individuals under anxiety prefer to stay cautious, in comparison with the neutral control group. The results show that negative affects can trigger different behaviors, thereby stating the failure of the valence-based approach.² This conclusion is endorsed by Fessler et al. (2004), who also raise the question of gender effect by demonstrating that anger increases risk-taking in men and unpleasantness reduces risk-taking in women.

An asymmetric influence of emotions is detected by Yuen and Lee (2003), in the sense that sadness leads to conservative decisions but no differences emerge between happiness and the neutral control. Chou et al. (2007) replicate the study by adding individual age as further treatment variable of the design. In this way, the authors find similar effect of sadness as to young participants, in addition to risk-seeking behavior under happiness when older subjects are compared. Basically, the AIM is fully supported only with reference to aged participants. A different story is told by Stanton et al. (2014), according to which the effect is asymmetric the other way round, with happy individuals being relatively risk-seeking and no differences between sad subjects and neutral ones.

The study by Zhao (2006) provides evidence in favor of the MMH, whereas Webb et al. (2012) conclude that individuals are willing to take more risks under unpleasantness, which still can be strategically controlled by deciding

¹The complete list of included papers is provided in Appendix 2A.

²Valence is the positive vs negative value of affect, nevertheless represents just one of the emotional factors influencing risky decision making (Lerner et al., 2015).

in advance how to act, that is, through implementation intentions.

Yip and Côté (2013) discover that incidental anxiety is able to reduce risk-taking, but the influence is counteracted in case one succeeds in identifying the source of her own emotion. Mixed evidence for the two main theories is offered by Drichoutis and Nayga Jr (2013) and Colasante et al. (2017), whose experiments elicit risk preferences through multiple price list formats and demonstrate that both positive and negative emotions lead to an increase in risk aversion.

Conte et al. (2018) also provide mixed endorsement, but in their case both positive and negative affects reduce cautiousness. Same findings are reported by Hu et al. (2015), who still fully endorse the AIM in case the experimental subjects make decisions under time pressure.

Campos-Vazquez and Cuijty (2014) investigate decision making under anger, sadness and fear, and conclude that sadness causes risk aversion but the other two affective states have no impact. Regarding sadness, similar yet weaker evidence is found by Treffers et al. (2016), whose results are also robust to variation of monetary incentives.

Finally, within a recent body of research aiming to give an emotion-based explanation for financial choices, Guiso et al. (2018) come to the conclusion that individuals under fear are willing to take less risks than the neutral control group. The study by Lee and Andrade (2015) confirms this result related to the financial context, as well as shows that the effect can be reversed by manipulating the framing of the decision task in such a way that fear is reinterpreted as excitement.

Such variety of findings is accompanied by further mixture of methods adopted, which makes it hard to explain the inconsistencies without resorting to quantitative techniques. Possible causes are discussed in the following section along with the coding of variables.

2.3 Data and methodology

In this section I offer an extensive view of the methodology adopted to calculate comparable effect sizes and select pertinent studies, before presenting the set of moderator variables for the meta-regression.

2.3.1 Effect size and identification of studies

The prototypical study that I aim to include in the analysis is based on a lab experiment consisting of treatments which differ only in the emotion artificially induced, although further control variables are free to change in the presence of factorial designs. After a manipulation check analyzes self-reported emotions and confirms the validity of the procedure in comparison with a neutral control, the subjects are supposed to carry out a task capturing risk-seeking behavior. At this point, any discrepancy in risk preferences between treatments can be ascribed to the different emotion experienced at the time of decision making and represents the effect size, that is, the summary statistic acting as dependent variable in the meta-regression. Since papers generally measure risk-seeking behavior through different scales and units of measurement, effect size usually takes the form of a standardized mean difference (SMD) which, following Cohen (1988)'s approach, is also known as Cohen's d :

$$SMD = d = \frac{\bar{Y}_E - \bar{Y}_C}{\sigma_{pooled}} \quad (2.1)$$

where \bar{Y}_E is the average outcome of the experimental group, \bar{Y}_C is the average outcome of the control group, and σ_{pooled} is the pooled standard deviation worked out like this:

$$\sigma_{pooled} = \sqrt{\frac{(n_E - 1)\sigma_E^2 + (n_C - 1)\sigma_C^2}{n_E + n_C - 2}} \quad (2.2)$$

with the sample sizes (n) also instrumental in obtaining an approximation of the standard error of the SMD (SE_{SMD}):

$$SE_{SMD} = \sqrt{\frac{(n_E + n_C)}{n_E * n_C} + \frac{SMD^2}{2 * (n_E + n_C)}} \quad (2.3)$$

However, researchers sometimes investigate the impact of emotions in the context of regressions, generally by adding a treatment dummy among the explanatory variables for each emotional group. In this case, as effect size I consider the related regression coefficient divided by the pooled standard deviation, which on this occasion utilizes the standard deviation of the treat-

ment dummy and the residual standard deviation in the computation. The procedure is endorsed by Chinn (2000), who also validates an approximate method to convert an odds ratio (OR) into SMD whenever the outcome variable is not continuous:

$$SMD = \frac{\sqrt{3}}{\pi} \ln(OR) \quad (2.4)$$

where the SMD is roughly equivalent to dividing $\ln(OR)$ by 1.81.³ Therefore, each effect size is a comparison of an experimental group under a certain emotion with a control group in a neutral condition, all the other variables being fixed.⁴ In case its value is positive (negative), then the subjects under the considered emotion made relatively risky (cautious) choices in relation to the neutral control group. An effect size close to zero denotes little influence of the corresponding emotion on risky decision making.⁵

Anyway, the identification of studies pertinent to the meta-analysis turns out to be rather cumbersome, since the sort of experiment being subject matter of the research does not own an universal denomination to be translated into keywords for a formal search, differently from other games in Experimental Economics. Moreover, the choice of appropriate keywords is hampered by a couple of motivations. First and foremost, a few manuscripts adopt a definition to conceptualize and another to operationalize, for instance by speaking about moods but at the same time employing stimuli to induce affective states, a procedure which according to the literature generates emotions (Robbins and Judge, 2012). Hence, I decide to consider the operational

³In the same way, $SE_{SMD} = \frac{SE_{\ln(OR)}}{1.81}$.

⁴For instance, Conte et al. (2018) run four treatments inducing different emotions (happiness, sadness, fear, anger) in addition to a neutral control treatment, this study thereby producing four observations in the meta-regression. Furthermore, scholars sometimes provide more than one estimate of the same effect by using different measures and techniques. Not to give excessive weight to a single study, in these cases I “choose the estimate that the author believes to be the best” (Stanley, 2001). When not possible, I pick the measure which is chronologically closer to the induction procedure, since it is a well-known fact that emotions are fleeting (Robbins and Judge, 2012). Finally, whenever an experiment is based on a factorial design or a paper is made up of more than one experiment, then I duplicate the number of observations accordingly. For further details, I refer the reader to Appendix 2B.

⁵Of note, in case a study analyzes risk-aversion instead of risk-seeking behavior, then I multiply its effect size by -1.

definition. Then, it is not infrequent that different terms are employed in order to refer to presumably the same emotion, both between and within articles.

Therefore, in an attempt to be inclusive I use Advanced Search in Google Scholar by looking for papers, even not published, which in the title contain either the word *risk* or *risky*, as well as one of the following terms: *emotion*, *emotions*, *mood*, *moods*.⁶ Furthermore, the word *experiment* needs to be present within the text, in order to easily filter out different methodologies. The resulting code is the following:

```
intitle:"risk OR risky" intitle:"emotions OR emotion OR moods OR mood" intext:experiment
```

with this search being conducted in July 2018 and yielding a list of 236 items.

At this point, for the sake of comparability the following criteria are applied:

1. the manuscript has to be grounded in a lab experiment, requirement which guarantees the identification of the causal nexus;
2. the experiment is between-subjects with respect to emotion, since the mixture with within-subjects studies is questionable from a statistical viewpoint (Morris and DeShon, 2002);
3. the experimental design needs to include an affect induction procedure, whose effectiveness is checked at least in a pre-test.⁷ In this way I can be sure that only specific emotions and not vague moods are analyzed, these emotions actually being present;

⁶Unfortunately, the asterisk wildcard character is not available in the Advanced Search environment.

⁷Among the studies included only Campos-Vazquez and Cuilty (2014) avoid testing the efficacy of the emotion induction procedure, simply because in their design the treatment variable is the intensity of the procedure and the subjects self-select into the emotion experienced.

4. a neutral control treatment is used and sufficient information needs to be available in the text so that the effect size can be computed through comparison.

This enables me to select 12 manuscripts, and other 4 are added after checking the references of the papers thus found.⁸ Finally, from the 16 studies which match the requirements I draw 46 observations at the treatment level, these articles being concentrated in the last two decades.

2.3.2 Moderator variables

Now I aim to choose a set of variables which represent experimental protocols and, according to related literature not investigating affects, may explain the inconsistencies in previous research findings.

Emotions

First of all, for each observation I need to control for the emotion induced in the treatment group, since different emotions might have dissimilar impact on risk-taking and, accordingly, on effect sizes. As mentioned in the previous subsection, this task is hindered by the fact that different terms are employed in order to refer to presumably the same emotion, both between and within articles. Therefore, having exploited information in the original texts, I seek to trace affects back to a common denominator in such a way as to solve the issue.

By so doing, in Table 2.1 for each umbrella emotion I am able to create a dummy variable equal to 1 in case the effect size originates from a comparison

⁸An example of a study not fulfilling the first criterion is Kliger and Levy (2003), where exogenous variations in weather are used as proxy for affective states. To give other examples, the design by Kuhnen and Knutson (2011) does not satisfy the second requirement, given that each subject experiences more than one emotion. Isen and Patrick (1983) does not comply with the third requisite, since no manipulation check is present in order to validate the emotion induction procedure. Instead, the fourth criterion is not satisfied in Heilman et al. (2010) due to the absence of a control treatment, which makes it impossible to calculate effect sizes. Anyway, a substantial number of papers are not considered because they investigate risk perception, which according to Weber and Milliman (1997) constitutes a different measure of risk attitude.

between that emotion and the neutral condition, and 0 otherwise.⁹

Table 2.1: Generating dummy variables for emotions

Umbrella emotion	Original terms
Happiness	Happiness; Joviality; Joy; Pleasure; Positivity
Sadness	Sadness; Grief
Unpleasantness	Unpleasantness; Disgust; Displeasure; Negativity
Anxiety	Anxiety
Fear	Fear; Fearfulness
Anger	Anger

Risk elicitation methods

At this point, because lots of different risk elicitation methods are used in the selected papers, I try to highlight homogeneous categories in an attempt to explain variation in effect sizes. Accordingly, a relevant distinction is that one between stated preferences and revealed preferences. Regarding the former, self-reported measures are definitely one of the simplest tools to elicit risk attitude, as described in Charness et al. (2013). In this case, each participant directly evaluates her own willingness to take risks in general or within specific domains, a non-incentivized procedure which still produces values correlated with the results of real-stakes lottery-based tasks (Dohmen et al., 2011). In addition, the absence of a task as filter makes it easier for an individual to appear risk-lover, which according to Camerer and Hogarth (1999) is a socially desirable attribute.

Instead, concerning revealed preferences there exists a vast array of more complex risk elicitation methods which would be arduous to consider as a whole. In line with the characteristics of the dataset, in this heterogeneous

⁹Especially in the case of “positivity” and “negativity”, which are maybe the most delicate terms to locate, I base my choice on similarities in the emotion induction procedure.

group I isolate the multiple price lists (MPL) à la Holt and Laury (2002), namely, a number of similar tasks where the switch point from a safer to a riskier option captures risk behavior in a multiple-choice environment. In their theoretical and experimental appraisal of four risk elicitation methods, Crosetto and Filippin (2016) point out that the MPL holds some features which do not favor risk-taking. Indeed, the final payment stems from a compound lottery instead of a simple one, a situation promoting risk aversion although the two cases should be equivalent under the Reduction Axiom (Kaivanto and Kroll, 2014). Moreover, the presence of a safe option in the list may induce certainty effects or act as a reference point.

Hence, in order to check if risk elicitation methods moderate the effect of emotions on risk attitude, I include a dummy *Stated preferences* equal to 1 if a self-reported measure of risk-taking is used, and 0 otherwise. The variable *MPL* is coded as 1 if risk preferences are elicited through a multiple price list, and 0 otherwise. Finally, a third dummy *Other task* takes on a value of 1 if revealed preferences are captured via other methods, mostly consisting in a single lottery choice. Given the related literature, I expect *MPL* to predict lower effect sizes than *Stated preferences*, the reference group.

Domain

Whether risk preferences are observed through self-reported measures (Hanoch et al., 2006) or elicited by means of more sophisticated tasks (Lee and Andrade, 2015), previous research demonstrates that context matters, in the sense that risk attitude is domain-specific and significantly depends on the framing of the elicitation method (Charness et al., 2013). In particular, Hanoch et al. (2006) notice that familiarity with the domain of choices lowers risk perception in such a way that, for instance, athletes take more recreational risks than subjects engaged in other professions.

In order to tackle the problem, the Domain-Specific Risk-Taking (DOSPRT) Scale (Blais and Weber, 2006) represents a valid solution via questionnaire, whereas lottery choices should be presented in an abstract fashion when it comes to measuring revealed preferences.

Anyway, many of the included studies analyze risk attitude under a single

domain of life (e.g., gamble, health, job security) which may moderate the influence of emotions. Among the various contexts, investment decisions are of particular interest to economists, therefore I single out the corresponding domain and assign a value of 1 to the dummy *Investment domain* when this framing is used. *Abstract domain* takes on a value of 1 if the context of the risk elicitation method is neutral and, finally, the variable *Other domain* is equal to 1 in case a different framing is employed.

Considering that (i) in a specific context (such as the financial one) a lower number of participants may feel at ease than in a generic abstract domain, and (ii) the experiments in the dataset mostly draw upon samples of students, then I expect *Investment domain* to be associated with lower effect sizes than *Abstract domain*, the omitted category.

Salient payment

Given the interdisciplinarity of papers in the meta-analysis, difference in experimental practices between Economics and Psychology is taken into account in the form of salient financial incentives, which are traditionally used only by economists. Indeed, there is solid evidence that monetary incentives reduce data variability and, when a standard for optimal behavior is present, “bring decisions closer to the predictions of the normative models” (Hertwig and Ortmann, 2001). Nevertheless, Camerer and Hogarth (1999) come to the conclusion that average behavior does not change to a great extent in tasks without a performance standard, such as choosing among gambles. The authors emphasize that only in a few circumstances related to risky decision making subjects take more risks when the payment is hypothetical, presumably due to the fact that risk-taking is a socially desirable quality. To see if this is also the case of the current dataset, I generate a dummy *Salient payment* equal to 1 in case the participants are remunerated according to their own actions, and 0 otherwise.¹⁰

¹⁰Random payment is regarded as salient, since it is well-established that this scheme can be equated with the pay-all approach (Cubitt et al., 1998; Hey and Lee, 2005). Instead, fixed payment is not considered to be salient (Friedman and Sunder, 1994).

Geographic variables

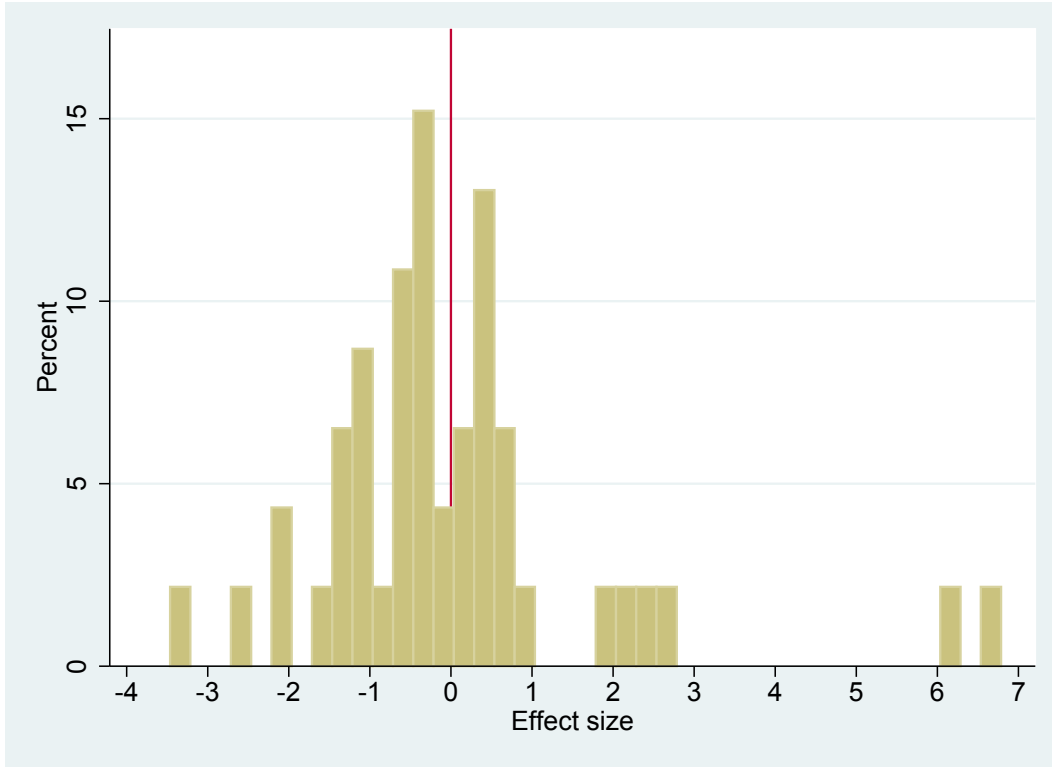
The dataset consists of 46 observations coming from experiments run in 4 continents: 18 data points from Europe, 13 from North America, 3 from Latin America, and 12 from Asia. As to the body of literature dealing with cultural differences in risk-taking, to date there is no agreement on the role played by local institutions. On the one hand, Hsee and Weber (1999) formulate the cushion hypothesis, according to which people from collectivist societies can rely on extended social networks able to give financial support, in case of need. This would lead to perceive the same uncertain choice as less risky and, accordingly, to take more risks than the citizens of individualistic societies, with this effect holding true only with respect to the investment domain. On the other hand, Kim and Park (2010) suggest that individualistic societies shape self-enhancing behavior and self-esteem, which are typically associated with risk propensity. In the absence of a predominant theory, the meta-regression might make a contribution to the debate.

2.4 Analysis and results

As said in Section 2.3, Cohen's d effect size is worked out for the 46 observations, this index being positive (negative) in case the experienced emotion leads to relatively risky (cautious) behavior, and close to zero when the effect is negligible.

The inconsistencies in previous findings are highlighted in Figure 2.1, which displays an overall effect size ranging from -3.5 to 6.7 and confirms the validity of the meta-analytic approach. Taking into account that this statistic does not distinguish the various emotions induced, the literature is characterized by effect sizes which differ in sign but are substantively of small magnitude, in line with the narrative review by Lane (2017). Indeed, the interval $[-0.5, 0.5]$ encompasses 41.3% of the observations. Moreover, a one-sample sign test is not able to reject the null hypothesis that the average effect size is equal to 0 ($p = 0.30$), which would comply with the mainstream theory predicting no influence of emotions on decision making. Egger's regression test excludes the presence of publication and related biases

Figure 2.1: Effect size distribution



($p = 0.49$).

At first sight, the high variation in effect sizes seems not to be explained even when I break down the dependent variable by emotion in Table 2.2. The insight is supported by the fact that across studies 5 of the 6 affective states have both positive and negative impact on risk-seeking behavior, with only *Anxiety* leading to outcomes of the same sign and low dispersion.

Furthermore, although in Figure 2.2 the distribution of the effect sizes related to *Anger* appears to be shifted towards greater risk, a Kruskal-Wallis test fails to reject the null hypothesis that the six emotion-based groups are drawn from the same population with respect to the median ($p = 0.36$). This preliminary result urges me to look at the other variables as possible sources of the study-to-study variation.

Hence, I estimate the following specification by using weighted least squares (WLS):

Table 2.2: Analyzing effect size

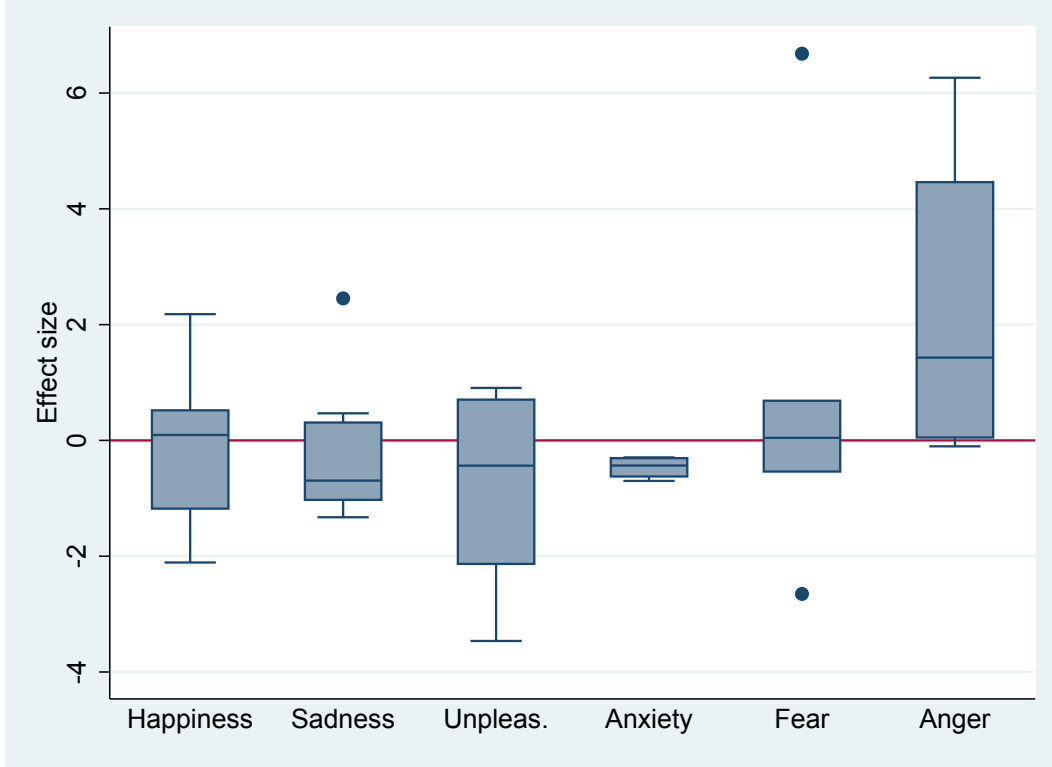
	Obs.	Effect size			
		Mean	Std. dev.	Min	Max
<i>Total</i>	46	0.016	1.844	-3.464	6.680
Happiness	12	-0.062	1.336	-2.109	2.180
Sadness	13	-0.358	1.060	-1.327	2.453
Unpleasantness	6	-0.810	1.692	-3.464	0.907
Anxiety	4	-0.465	0.193	-0.700	-0.293
Fear	7	0.548	2.903	-2.652	6.680
Anger	4	2.256	2.944	-0.103	6.264
Stated preferences	11	-0.141	0.957	-1.281	2.009
MPL	13	-0.919	1.005	-3.464	0.203
<i>Other task</i>	22	0.647	2.303	-2.652	6.680
Abstract domain	32	0.148	2.089	-3.464	6.680
Investment domain	2	-1.572	1.528	-2.652	-0.491
<i>Other domain</i>	12	-0.072	0.943	-1.281	2.009
Salient payment	19	0.458	2.617	-3.464	6.680
Non-salient payment	27	-0.295	0.950	-2.652	2.009
Asia	12	-0.072	0.983	-1.309	2.009
North America	13	-0.276	1.275	-2.652	2.658
Latin America	3	-0.035	0.332	-0.414	0.203
Europe	18	0.294	2.658	-3.464	6.680

$$y_i\sqrt{w_i} = \beta_1\sqrt{w_i} + \beta_2x_{i2}\sqrt{w_i} + \dots + \beta_kx_{ik}\sqrt{w_i} + \varepsilon_i\sqrt{w_i} \quad (2.5)$$

where y_i is the effect size, x_{ij} stands for the j th methodological variable out of k regressors, and w_i represents the analytic weight for the i th observation.

Indeed, the rationale behind the use of weights lies in the fact that effect sizes from larger studies are more precise than those from smaller studies, and accordingly should be given more weight. Following Hedges and Olkin (1985), each analytic weight is conceived in comparison to the sum of the 46 datapoints and computed in the ensuing fashion:

Figure 2.2: Effect size by emotion



$$w_i = \frac{1}{\hat{\sigma}_i^2} = \frac{n_i^E n_i^C / (n_i^E + n_i^C)}{\sum [n_i^E n_i^C / (n_i^E + n_i^C)]} \quad (2.6)$$

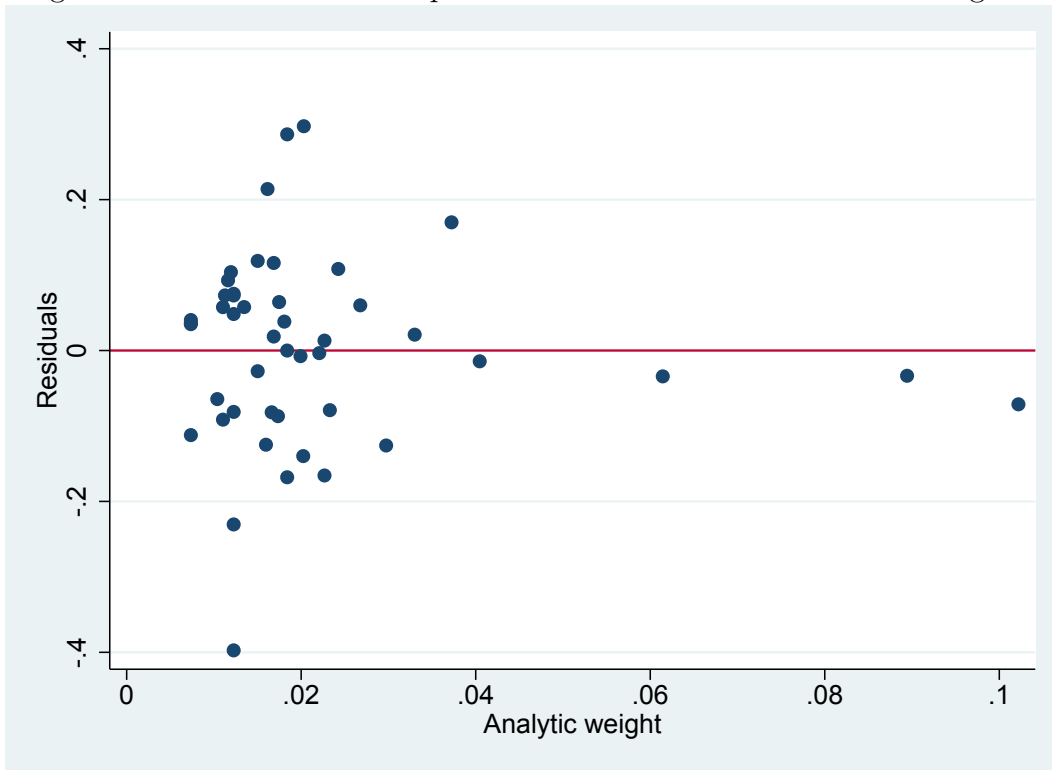
with n_i^E and n_i^C being the number of subjects in the experimental group and in the control group for the i th observation, respectively.¹¹

Moreover, the WLS estimator is deemed to be superior to conventional random-effects, mixed-effects and fixed-effects counterparts (Stanley and Doucouliagos, 2017), as well as an elegant approach to deal with the problem of heteroskedasticity, which is typical of cross sectional data. The choice of giving more weight to larger studies turns out to be methodologically sound, as

¹¹Given that sometimes the selected papers indicated only the total number of participants without making reference to n_i^E and n_i^C , I e-mailed all the authors in order to calculate the analytic weights (and the effect sizes as well). In the few cases where I received no answer, then I realistically assumed that the subjects were equally distributed between experimental and control group. The related information is provided in Appendix 2B.

depicted in Figure 2.3 which actually displays an inverse relationship between variance of residuals and the analytic weights.

Figure 2.3: Inverse relationship between variance of residuals and weights



At a later stage, I bear in mind that each manuscript is free to contribute more data points to the meta-regression, so that it is advisable to take action and allow for dependence between observations. Accordingly, I correct the results by clustering standard errors at the level of paper.

Taking into consideration the small number of observations, a further concern is represented by the fact that OLS normality assumption may be violated. Indeed, the distribution of effect sizes fails the skewness-kurtosis tests for non-normality, with $p < 0.01$ in both cases. Therefore, as additional robustness check I accompany the previous models with another one where the standard errors are worked out by means of the block bootstrap technique (Moffatt, 2015).¹²

¹²As to the number of bootstrap samples, Moffatt (2015) reports that three benchmarks

Table 2.3: Explaining effect size

	Dependent variable: Effect size (tr)		
	(1) WLS	(2) Clustered WLS	(3) B. Bootstrapped WLS
MPL (tr)	-5.881*** (1.480)	-5.881*** (1.287)	-5.881** (2.434)
<i>Other task (tr)</i>	-1.850 (1.375)	-1.850 (1.161)	-1.850 (1.289)
Investment domain (tr)	-3.552*** (0.840)	-3.552*** (0.828)	-3.552*** (1.237)
<i>Other domain (tr)</i>	-1.890 (1.357)	-1.890* (1.031)	-1.890 (1.256)
Salient payment (tr)	0.955* (0.484)	0.955* (0.489)	0.955 (0.827)
North America (tr)	-2.601*** (0.596)	-2.601*** (0.542)	-2.601* (1.483)
Latin America (tr)	0.721 (0.716)	0.721 (0.594)	0.721 (0.957)
Asia (tr)	-2.019*** (0.686)	-2.019*** (0.587)	-2.019 (1.309)
Sadness (tr)	0.060 (0.472)	0.060 (0.639)	0.060 (0.544)
Unpleasantness (tr)	-1.708** (0.680)	-1.708** (0.654)	-1.708* (1.031)
Anxiety (tr)	0.209 (0.773)	0.209 (0.733)	0.209 (0.914)
Fear (tr)	1.683** (0.611)	1.683 (1.068)	1.683* (0.908)
Anger (tr)	1.669** (0.660)	1.669 (0.983)	1.669* (0.859)
$\sqrt{w_i}$	3.805*** (1.338)	3.805*** (0.964)	3.805** (1.783)
Prob > F	0.000	-	-
R-squared	0.763	0.763	0.763
Adj R-squared	0.659	-	0.659
Observations	46	46	46

WLS coefficient estimates with standard errors in parentheses, which are clustered at the level of paper in column (2) and worked out with block bootstrap technique in column (3). (*tr*) indicates the variable transformed by the square root of the analytic weight, which consequently becomes the constant of the models. The omitted categories are *Stated preferences*, *Abstract domain*, *Non-salient payment*, *Europe* and *Happiness*.

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

Table 2.3 suggests that the issues related to dependence between observations and the violation of OLS normality assumption may not be serious, since the three models produce similar results.

After controlling for the specific type of emotion investigated, some experimental protocols appear to moderate the relationship between affect and risk in a way that is coherent with related literature. In particular, studies using a multiple price list à la Holt and Laury (2002) observe significantly lower effect sizes than manuscripts collecting self-reported measures of risk attitude ($p < 0.01$). As remarked in Section 2.3, Crosetto and Filippin (2016) already speculated that some aforementioned characteristics of this format foster risk aversion, when comparing tasks that gauge revealed preferences. Actually, the insight still holds true under emotions and in relation to stated preferences methods.

Likewise, I find significant evidence that emotions induce greater risk aversion when the risk elicitation method is framed as an investment decision in place of an abstract choice ($p < 0.01$), this outcome supporting the idea that reduced familiarity with the context is associated with higher risk perception (Hanoch et al., 2006). Instead, remunerating participants in a salient way seems not to bias effect sizes but increases their magnitude (Levene's test, $p < 0.01$), at odds with evidence in Hertwig and Ortmann (2001).

While all these findings are robust to the different techniques for computing standard errors, the effects for the geographic regions are more mixed. When participating in experiments conducted in North America, subjects take less risks than in Europe (the omitted category), this being the only geographic variable which preserves significance in all three columns. This result is also hard to be explained through any culture-based theory, hinting at the role of other unobserved characteristics of these regions.

Finally, Table 2.4 varies the omitted category and reports the coefficients for the dummies controlling for the emotions, when standard errors are calculated with block bootstrap technique. In this way the reader can appreciate that actually the six affective states lead to different outcomes with respect

appear to have emerged, that is, either 99 or 999 or 9999. I go for the last-mentioned value, since a substantial number of replications are dropped due to perfect multicollinearity.

Table 2.4: Different impact of emotions

	Omitted category					
	Happiness	Sadness	Unpleas.	Anxiety	Fear	Anger
Happiness	-	-0.060	1.708	-0.209	-1.683*	-1.669*
Sadness	0.060	-	1.768	-0.149	-1.623***	-1.609***
Unpleas.	-1.708*	-1.768	-	-1.917	-3.391**	-3.377**
Anxiety	0.209	0.149	1.917	-	-1.474*	-1.460*
Fear	1.683*	1.623***	3.391**	1.474*	-	0.014
Anger	1.669*	1.609***	3.377**	1.460*	-0.014	-

WLS coefficient estimates with standard errors worked out with block bootstrap technique, by varying the omitted category.

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

to each other. This remark emphasizes the failure of the valence-based approach, given that emotions of the same valence do not have similar effects on risk attitude. Specific affects should be chosen when developing the research question.

2.5 Conclusions

The chapter is a meta-analysis of experimental studies dealing with the impact of incidental emotions on risky choices, so as to explain traditional heterogeneity of outcomes in the field. In the text I accepted the invitation from Lane (2017), who had previously reviewed the area of interest by underlining the necessity of a meta-regression to account for inconclusive findings.

After performing a weighted least squares (WLS) regression where larger studies are given more weight, I found that emotions induce higher risk aversion when a multiple price list à la Holt and Laury (2002) is used in place of stated preferences methods, as well as in case the risk elicitation task is framed as an investment decision instead of an abstract choice. These results were robust to different techniques for computing standard errors and in line with related literature not inspecting affects. The regressions also showed that stakes do not systematically influence behavior yet salient payments in-

crease variability of outcomes, thereby contradicting solid evidence (Hertwig and Ortmann, 2001). Moreover, the role of geographic differences remained unclear and further investigation of this aspect might be a fruitful endeavor. Overall, anxiety turned out to be the only emotion leading to effects of the same sign, in the direction of greater risk aversion. At the same time, the data confirmed that the valence-based approach should not be followed, since emotions of the same valence trigger significantly different risky choices.

To conclude, regarding the limitations of the current study, I recognize that gender, age and other demographics might play a part in shaping risk preferences. Nevertheless, it is cumbersome to investigate this at the level of treatment. All in all, given the array of protocols used in this kind of experiments and in the absence of a tailor-made game to answer such research questions, I recommend faithful study replication as preferential path in order to analyze the impact of emotions on risky decision making and ensure comparability.

Appendix 2A. Papers included in meta-analysis

1. **Campos-Vazquez, R.M. and Cuijly, E. (2014).** The role of emotions on risk aversion: A Prospect Theory experiment. *Journal of Behavioral and Experimental Economics*, 50: 1-9.
2. **Chou, K.L., Lee T., and Ho, A.H. (2007).** Does mood state change risk taking tendency in older adults? *Psychology and Aging*, 22(2): 310-318.
3. **Colasante A., Marini, M.M., and Russo, A. (2018).** Incidental emotions and risk-taking: An experimental analysis. *Working Papers 2018/13, Economics Department, Universitat Jaume I, Castellón (Spain)*.
4. **Conte, A., Levati, M.V., and Nardi, C. (2018).** Risk preferences and the role of emotions. *Economica*, 85(338): 305-328.
5. **Drichoutis, A.C. and Nayga Jr, R.M. (2013).** Eliciting risk and time preferences under induced mood states. *The Journal of Socio-Economics*, 45: 18-27.
6. **Fessler, D.M., Pillsworth, E.G., and Flamson, T.J. (2004).** Angry men and disgusted women: An evolutionary approach to the influence of emotions on risk taking. *Organizational Behavior and Human Decision Processes*, 95(1): 107-123.
7. **Guiso, L., Sapienza, P., and Zingales, L. (2018).** Time varying risk aversion. *Journal of Financial Economics*, 128(3): 403-421.
8. **Hu, Y., Wang, D., Pang, K., Xu, G., and Guo, J. (2015).** The effect of emotion and time pressure on risk decision-making. *Journal of Risk Research*, 18(5): 637-650.
9. **Lee, C.J. and Andrade, E.B. (2015).** Fear, excitement, and financial risk-taking. *Cognition and Emotion*, 29(1): 178-187.
10. **Raghunathan, R. and Pham, M.T. (1999).** All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes*, 79(1):56-77.
11. **Stanton, S.J., Reeck, C., Huettel, S.A., and LaBar, K. S. (2014).** Effects of induced moods on economic choices. *Judgment and Decision Making*, 9(2): 167-175.
12. **Treppers, T., Koellinger, P.D., and Picot, A. (2016).** Do affective states influence risk preferences? *Schmalenbach Business Review*, 17(3-4): 309-335.
13. **Webb, T. L., Sheeran, P., Totterdell, P., Miles, E., Mansell, W., and Baker S. (2012).** Using implementation intentions to overcome the effect of mood on risky behaviour. *British Journal of Social Psychology*, 51(2): 330-345.

14. **Yip, J.A. and Côté, S. (2013).** The emotionally intelligent decision maker: Emotion-understanding ability reduces the effect of incidental anxiety on risk taking. *Psychological Science*, 24(1): 48-55.
15. **Yuen, K. S. and Lee, T.M. (2003).** Could mood state affect risk-taking decisions? *Journal of Affective Disorders*, 75(1):11-18.
16. **Zhao, J. (2006).** The effects of induced positive and negative emotions on risky decision making. *Talk presented at the 28th Annual Psychological Society of Ireland Congress, Maynooth, Ireland.*

Appendix 2B. Data

Paper	Emotion treated	Region	Sample size (treatment + control)	Effect size (SMD)	Stated prefer.	MPL	Other task	Abs. domain	Investm. domain	Other domain	Salient paym.
Campos-Vazquez and Cuijly (2014)	Anger	Latin Am.	336*	0.20		✓		✓			
	Sadness	Latin Am.	264*	-0.41		✓		✓			
	Fear	Latin Am.	360*	0.11		✓		✓			
Chou et al. (2007)	Happiness	Asia	65	-0.16	✓					✓	
	Sadness	Asia	66	-1.03	✓					✓	
	Happiness	Asia	60	2.01	✓					✓	
	Sadness	Asia	60	-1.28	✓					✓	
Colasante et al. (2018)	Happiness	Europe	88	-0.43		✓		✓			✓
	Sadness	Europe	93	-0.69		✓		✓			✓
Conte et al. (2018)	Happiness	Europe	97	2.18				✓			✓
	Sadness	Europe	80	2.45				✓			✓
	Fear	Europe	74	6.68				✓			✓
	Anger	Europe	67	6.26				✓			✓
Drichoutis and Nayga Jr (2013)	Happiness	Europe	53*	-2.11		✓		✓			✓
	Unpleas.	Europe	56*	-3.46		✓		✓			✓

Appendix 2B (continued)

Paper	Emotion treated	Region	Sample size (treatment + control)	Effect size (SMD)	Stated prefer.	MPL	Other task	Abs. domain	Investm. domain	Other domain	Salient paym.
Fessler et al. (2004)	Anger	North Am.	40	2.66			✓	✓			✓
	Unpleas.	North Am.	39	-0.45			✓	✓			✓
	Anger	North Am.	40	-0.10			✓	✓			✓
	Unpleas.	North Am.	40	-2.13			✓	✓			✓
Guiso et al. (2018)	Fear	North Am.	207	-2.65			✓		✓		
	Happiness	Asia	24**	0.34			✓	✓			
Hu et al. (2015)	Sadness	Asia	24**	0.47			✓	✓			
	Happiness	Asia	24**	0.36			✓	✓			
	Sadness	Asia	24**	-1.31			✓	✓			
	Fear	Asia	57	-0.49			✓		✓		✓
Lee and Andrade (2015)	Fear	Asia	60	0.69			✓			✓	✓
Raghunathan and Pham (1999)	Anxiety	North Am.	55**	-0.29			✓	✓		✓	
	Sadness	North Am.	55**	0.31			✓	✓			
	Anxiety	North Am.	49**	-0.70	✓					✓	
	Sadness	North Am.	49**	0.34	✓					✓	
Stanton et al. (2014)	Happiness	North Am.	60*	0.59*			✓	✓			✓
	Sadness	North Am.	57*	-0.29*			✓	✓			✓
Treffers et al. (2016)	Happiness	Europe	38	-1.21*		✓		✓			
	Fear	Europe	40	0.04*		✓		✓			
	Sadness	Europe	37	-1.33*		✓		✓			
	Happiness	Europe	72	-1.14*		✓		✓			✓
	Fear	Europe	74	-0.54*		✓		✓			✓
	Sadness	Europe	74	-0.97*		✓		✓			✓

Appendix 2B (continued)

Paper	Emotion treated	Region	Sample size (treatment + control)	Effect size (SMD)	Stated prefer.	MPL	Other task	Abs. domain	Investm. domain	Other domain	Salient paym.
Webb et al. (2012)	Unpleas.	Europe	44*	0.70	✓						✓
	Unpleas.	Europe	34*	-0.42	✓						✓
Yip and Côté (2013)	Anxiety	North Am.	108*	-0.32			✓				
	Anxiety	North Am.	132*	-0.55	✓						✓
Yuen and Lee (2003)	Happiness	Asia	36**	0.44	✓						✓
	Sadness	Asia	36**	-0.91	✓						✓
Zhao (2006)	Happiness	Europe	40	-1.63			✓				✓
	Unpleas.	Europe	40	0.91			✓				✓

**Subjects are assumed to be equally distributed between treatment and control group.

*Elaborated from the raw data.

Chapter 3

Communication in a threshold public goods game with ambiguity: Anomalies and regularities

3.1 Introduction

Decision makers often end up with private information as the only estimate of common-value public goods whose profitability is pervaded by uncertainty (Cox, 2015). For instance, investment in prevention typically produces variable outcomes due to the unpredictable nature of natural disasters and epidemics, in such a way that even the investors themselves cast doubt on the efficacy of such actions. In these environments, should the members of a community be able to aggregate information, uncertainty would definitely be reduced.

In an attempt to consider concrete applications, let me focus on joint research projects as a case where the quality of the final output cannot be inferred in advance. Indeed, the decision to join with colleagues to write a manuscript can be seen as a binary contribution to a public good which is provided only in case a threshold number of coauthors is reached, each one possessing skills needed for the successful completion. In support of this view,

also the researchers of the same university department who are not involved in a certain project can benefit from the publication of a paper, in terms of increased reputation and attractiveness of the institution. Clearly, authors and non-authors do not obtain the same returns on this public good, which still are ambiguous and could be better predicted by means of communication. Repeated interactions help to minimize uncertainty by offering a glimpse of how a possible collaboration might be: if my colleagues are not very responsive to e-mails or fall short of my expectations, maybe I had better work by myself. But in case information aggregation raises expectations to an extent that the production of a paper with actual low impact leaves coauthors disappointed, future willingness to collaborate might be impaired. Basically, low quality of the public good could hamper successive contributions. In support of this argument, disappointment originates from the comparison between a current event and a better state of the world which has not materialized, taking the shape of “a psychological reaction to an outcome that does not match up to expectations” (Bell, 1985). Moreover, undertaking a joint project without the overall necessary skills is definitely harmful, but having a high number of coauthors may reveal minor inefficiency as well, thereby leading to the final output at excessive cost in terms of prolonged time before submissions. Therefore, it would be interesting to test if the opportunity to communicate actually yields efficient outcomes also under uncertainty about the value of the public good, given this gap in the literature on cooperation. This chapter intends to replicate the whole process in a controlled setting, so as to detect possible anomalies.

In order to investigate cooperative behavior, I run a laboratory experiment with 160 participants equally split into two treatments, where the control variable is pre-play communication in the form of unrestricted text chat. A binary threshold public goods game with four-person groups and a threshold of three is at the core of the design, which also includes a few secondary tasks measuring variables of interest. The main novelty I introduce is related to the provision mechanism which, in case the threshold is reached, consists in a binary lottery producing either a stated high value or a stated low value for the whole group with fixed unknown probability over the fifteen periods played.

In general, during each round of the game the players are supposed to coordinate on two levels: (i) the decision about whether to reach the threshold and, if so, (ii) the decision about how many group members should contribute. Before the contribution decision, in each period decentralized information is provided by means of a private signal for the actual value of the public good, with this signal being dichotomized for the sake of interpretation. Concerning the two levels, (i) is related to the players' beliefs on the binary lottery generating the public good and, accordingly, relies on the private signal. Instead, as to (ii) no hint is offered and, in the absence of communication, the most efficient outcome can occur only by chance. On the contrary, the group members in the chat treatment can resort to a valid coordination device in order to reach the agreement on three contributors. In this case, the partner matching allows the groups to implement a rotation of the only one member not contributing.

Following Palfrey et al. (2017)'s approach, I show that my game involves three sources of inefficiency: (a) undercontribution, in case just one or two group members contribute and the threshold is not achieved; (b) overcontribution, when all four players contribute and the public good is provided at excessive cost; (c) the usual free-rider problem.

The results at the group level emphasize that, in accordance with related literature, communication significantly increases public good provision by reducing inefficiency that comes from wasteful undercontribution. The observed higher provision rate in the chat treatment is fostered by the fact that the group members hardly ever decide to coordinate on zero contributions when they are free to communicate. Indeed, contrary to my hypothesis based on Bell (1985)'s theory of disappointment in decision making under uncertainty, the players are not let down when they run into low quality of the public good, after reaching the threshold. Moreover, the subjects in the chat treatment seem to neglect the free-rider issue and often end up overcontributing, in contrast with previous scientific findings.

After chat analysis, I discover that the issue of overcontribution is even more pervasive than what can be observed from the mere number of group contributors, due to moderate presence of defectors within agreements on all members contributing. Nevertheless, the deals on this action profile are

often successful, arguably thanks to subjects' willingness to pursue symmetric payoffs as well as to pronounced group identity generated by the partner matching and the common fate hypothesis.

To conclude, despite the presence of a valid coordination device as the rotation of the non-contributor, the group members privilege the implementation of strategies based on the systematic disclosure of all four private signals. Since the players prefer to minimize ambiguity than to maximize the group earnings, I finally speculate that under uncertainty satisficing is more salient than optimizing.

The remainder of the chapter is organized as follows. Section 3.2 reviews the related literature. Section 3.3 includes a detailed description of the experimental phases as well as the hypotheses of the present work. The results are presented in Section 3.4, where after carrying out an overall analysis, I test the hypotheses at the group level and propose possible explanations. Section 3.5 concludes.

3.2 Literature review

Prominent surveys on public goods games (Ledyard, 1995; Chaudhuri, 2011; Dhami, 2016) identify communication as one of the non-punitive methods in order to sustain cooperation. Although face-to-face communication turns out to be the most effective form, still communication via computer monitor outperforms the no communication condition on promoting cooperation (Bochet et al., 2006). After reviewing the experimental literature on social dilemmas, Bicchieri and Lev-On (2007) theorize the positive influence of the communication effect also in computer-mediated environments, by stressing that the effect is as powerful as these environments are able to reproduce the features of face-to-face communication. Palfrey et al. (2017) add that the effect of pre-play communication on efficiency depends on the richness of the message space, since the efficient outcomes observed under unrestricted text chat are not found when binary or numerical messages are allowed. From a theoretical viewpoint, Agastya et al. (2007) indicate that binary communication can provide individuals with incentives to contribute, with respect to absence of communication. As far as coordination games are concerned, the picture

becomes more complicated: one-way communication should be preferred in games of conflict like the battle of the sexes, whereas two-way communication is more effective when there is no dominated strategy (Cooper et al., 1992). Therefore, non-binding pre-play communication does not always entail efficiency gains, which still occur in case players decide to disclose private information (Crawford and Sobel, 1982).

Traditionally, a number of moderators whereby communication operates have been detected, such as the development of group identity, the change of expectations of others' behavior and the offering of promises (Ostrom et al., 1994). Nevertheless, even deviations from pledges on Nash equilibria are observed (Tavoni et al., 2011).

However, the current work can be framed in the branch of the so-called step-level public goods (SLPG) games, where the public good value is provided as long as a certain threshold is overcome.¹ Among the most common examples, the construction of infrastructures, the passing of a law in parliament or fundraising projects are just some cases in which the quality or quantity of the public good do not increase significantly if the contributions exceed the threshold (Offerman et al., 1996). As Alberti and Cartwright (2016) point out, a relevant difference from standard public goods is that there is not the typical conflict between the maximum individual payoff and the best social outcome, since some Nash equilibria coincide with the most efficient combination of strategies. This implies that, in addition to cooperation, the current problem becomes a matter of coordination. In the last decade, SLPG games have been repeatedly employed within the context of environmental issues, notably climate change, where each individual is supposed to make an effort so as not to surpass thresholds of environmental resources whose excessive use could lead to harmful consequences (Tavoni et al., 2011; Barrett and Dannenberg, 2012; Güth et al., 2015). In this frame, elements of ambiguity are often introduced to simulate uncertainty with respect to the level of the threshold, which usually is unknown.

Ambiguity, also known as uncertainty, refers to the case where the outcome probabilities of a risky choice are not known by the decision makers.

¹Indeed, these public goods are also called threshold public goods, or public goods with a provision point mechanism (PPM).

On the one hand, uncertainty has been studied long in the field of individual decision making after Ellsberg (1961)'s criticism of Savage's expected utility theory with subjective probabilities, mainly by postulating individual ambiguity aversion. On the other hand, the interest in delving into ambiguity within strategic environments, where ambiguity is usually elaborated as uncertainty about the behavior of the others, has arisen only recently. Following this approach, evidence that ambiguity is able to blunt the free-riding problem has proliferated in the last decades (Eichberger and Kelsey, 2002; Bailey et al., 2005; Keenan et al., 2006). Moreover, safe strategies appear more attractive under uncertainty, which is consistent with the ambiguity aversion hypothesis (Kelsey and Le Roux, 2017; Calford, 2017).

Anyway, my study departs from the experimental literature as to a set of aspects. The main novelty is related to the way ambiguity is introduced in the experimental design: if the threshold is reached, the provision mechanism consists in a binary lottery producing either a stated high value or a stated low value for the whole group with fixed unknown probability over periods. Otherwise, the contributions are lost and the lottery is not played. Moreover, each player is provided with a private signal for the actual value of the public good so that coordination might be hampered, given the heterogeneity of such signals (Güth et al., 2015) and the free-riders' incentive to lie and overstate the public good value (Crawford and Sobel, 1982). Another seminal feature is that each group member is allowed to contribute either a fixed amount or nothing, in compliance with the tradition of binary SLPG games where a threshold number of contributors are needed for the production of the public good (Dawes et al., 1986; Palfrey and Rosenthal, 1991; Offerman et al., 1998). As Offerman et al. (1996) emphasize, in this instance a crucial role is played by the individual belief about what the other group members do, since everyone would like her own contribution to be critical for the threshold achievement. At the same time, the authors notice that the contribution decision might be driven by social orientation, namely, the weight that one ascribes to her own welfare with respect to the others' well-being. These factors may interact with the determinants of cooperation which are object of study: therefore, I am going to elicit and control for them through incentive-compatible mechanisms, as described in Section 3.3.

All the aforementioned characteristics are fixed throughout the experiment, whose control variable is the presence of pre-play communication in the form of unrestricted text chat.

3.3 Experimental design and hypotheses

The current section includes a detailed description of the experimental phases as well as the hypotheses I am going to test.

A threshold public goods game is at the core of the design, which also encompasses a Ring-test capturing social orientation and a few questionnaires measuring variables of interest.

Table 3.1: Structure of the experiment

Treatment	N° of sessions	Subjects	Females
No communication (NC)	2	80	50%
Chat	2	80	41%

As illustrated in Table 3.1 and Table 3.2, I run two treatments constituted by two sessions each. The treatment variable is contained in the public goods game and is represented by unbinding pre-play communication in the form of unrestricted text chat. Throughout the experiment salient financial incentives are provided, since one period of the public goods game is randomly chosen and a further lottery between that period and the Ring-test outcome determines the final payment, in addition to a show-up fee of 5 euro. The control treatment takes one hour and a half to be carried out. I increase the show-up fee to 7 euro in the chat sessions, which last two hours. The experimental currency is the ECU (Experimental Currency Unit), with the exchange rate $5 \text{ ECU} = 1 \text{ euro}$ in both the remunerated tasks of the experiment. In the end, the subjects in the baseline gain on average 18.4 euro, whereas the mean earnings in the chat treatment amount to 20.5 euro.

The experiment is programmed by using the software z-Tree (Fischbacher, 2007) and takes place at the Laboratorio de Economía Experimental (LEE)

Table 3.2: Structure of each session

Order	Phase
1	Control questions
2	Threshold public goods game with(out) chat
3	Socio-demographic questionnaire
4	Ring-test
5	Interdependency and empathy questions
6	Payment feedback

of Universitat Jaume I (Castellón, Spain). Overall, I recruit 160 students (73 females) from the university campus through ORSEE (Greiner, 2015) and equally split them into four sessions.

3.3.1 The threshold public goods game

In each session, after the roll call the subjects are randomly allocated to the cubicles and individually read the instructions of the first task. Before starting the public goods game, the participants are supposed to answer five control questions by which I want to test their grasp of the rules and incentives of the game. Once the solutions are explained aloud, the game begins and four-person groups are formed with partner matching over fifteen periods. At the beginning of every period, each group member is endowed with 50 ECU and can either contribute the entire endowment to a public good or decide not to contribute by bearing a smaller cost equal to 15 ECU.²

The main novelty I introduce is related to the provision mechanism which consists in a binary lottery producing either a high value of 80 ECU or a low value of 20 ECU for each group member.

Such a lottery is based on a box containing 100 balls, and each of them holds the same unknown probability of being white. This probability is set

²I introduce a cost for not contributing because the decision of not joining with colleagues in the department to write manuscripts entails a potential harm to oneself, and not only to the other members of the community.

to 70% and concealed from the subjects, who only know that the probability is constant throughout the treatment. Every period, in case the threshold of three contributions out of four is reached, one ball is drawn: if the ball is white, each member wins 80 ECU. Otherwise, 20 ECU. In case the threshold is not achieved, the contributions are not refunded and the lottery is not played.

Moreover, before making the contribution decision, every player is provided with a private signal which, for the sake of interpretation, is dichotomized.³ Indeed, each group member draws an independent sample of 10 balls from the box and subsequently receives a message which can be of two kinds:

- the sample proportion of white balls is high (greater than or equal to 5/10);
- the sample proportion of white balls is low (less than 5/10).

Following Palfrey et al. (2017)'s framework, given a group made up of N individuals, the public good is yielded only if at least K members contribute. Accordingly, the return net of endowment for player i amounts to:

$$\begin{array}{ll}
 -15 + g & \text{if } i \text{ does not contribute and at least } K \text{ others contribute} \\
 -15 & \text{if } i \text{ does not contribute and fewer than } K \text{ others contribute} \\
 -50 + g & \text{if } i \text{ contributes and at least } K - 1 \text{ others contribute} \\
 -50 & \text{if } i \text{ contributes and fewer than } K - 1 \text{ others contribute}
 \end{array}$$

where clearly $N = 4$, $K = 3$ and $g \in \{20, 80\}$.

Therefore, the game involves three sources of inefficiency: (a) undercontribution, in case just one or two group members contribute and the threshold is not achieved; (b) overcontribution, when all four players contribute and the public good is provided at excessive cost; (c) the free-rider problem, since everyone would prefer to enjoy the public good, whether high-valued or low-valued, without contributing.

³From a programming viewpoint, in each period four independent samples are drawn and then replicated in all the groups of the session.

In the SLPG game I also address the issue of individual expectations, that is, I take into account the individual belief about what the other group members do. As in Caporael et al. (1989), a contribution can be categorized as *futile* if fewer than $K - 1$ others contribute, *critical* if exactly $K - 1$ others contribute or *redundant* if K or more others contribute. Since Offerman et al. (1996) notice a positive correlation between the subjective probability of being critical and the contribution choice, in my design I induce the participants to truthfully reveal their own subjective probability of being critical by means of an incentivized question, which is placed immediately after the contribution stage.⁴ Hence, in each period the individual profit is the algebraic sum of three components: the initial endowment, the net return and the gain coming from the incentivized question.

At the end of every period, the players are given information on their own individual profit, on the threshold achievement and on the outcome of the lottery generating the public good value by means of a history table, which summarizes relevant variables of all periods and stays visible in all stages of the treatment. No feedback is provided about how many group members contribute or about the period randomly selected for the final payment.

As said, the control variable is represented by unbinding pre-play communication in the form of unrestricted text chat. Indeed, having received the private signals, the players in the treatment sessions are allowed to send messages to the other group members for an interval of 90 seconds. Moreover, the subjects are warned that such messages need be relevant to the experiment and not intended to reveal their identity. Once time expires, the chat freezes but the messages exchanged remain visible until the participants make the contribution decision.

⁴In order to induce risk neutrality and obtain reliable answers, I combine the quadratic scoring rule with a binary lottery procedure, as described in Harrison et al. (2014). By so doing, the more accurate the prediction, the higher (lower) the chances of winning 5 extra ECU (nothing). Being θ one's subjective probability that exactly two out of the other three group members contribute, the probability of winning 5 extra ECU is equal to $P(\theta | C) = 100 - 100(1 - \theta/100)^2$ in case actually two out of the other three group members contribute. Otherwise, that probability is equal to $P(\theta | NC) = 100 - 100(0 - \theta/100)^2$.

I minimize the possible hedging problem by keeping the stakes for belief elicitation "small" relative to the other choice tasks, in compliance with the literature (Blanco et al., 2010).

Before moving to the following task, the participants fill in a socio-demographic questionnaire which also gathers data on self-reported risk aversion.

3.3.2 The Ring-test and the final questionnaires

At this point, I remind the subjects that the next task is unrelated to the first part of the experiment and I hand out the instructions of the Ring-test, also known as Decomposed Game (Liebrand, 1984), by which I want to control for social orientation. Afterwards, I explain aloud the rules and, once all doubts are clarified, the game starts.

In this instance, each player is randomly matched with another participant, whose identity remains unknown throughout the task. Each subject faces 24 scenarios and in each of them she is supposed to choose between two options, “A” and “B”. Each option assigns a positive or negative amount of ECU to the subject herself (payoff x) and to her partner (payoff y). In other words, the players are asked to pick their preferred money allocation between the two options displayed in each stage. Until the end of the session, no feedback is given about the partner’s choices and about the individual profit, which is calculated by adding up the amount the subject assigns to herself and the money she receives from the partner, in all scenarios. This algebraic sum is added to (subtracted from) an initial endowment of 40 ECU.

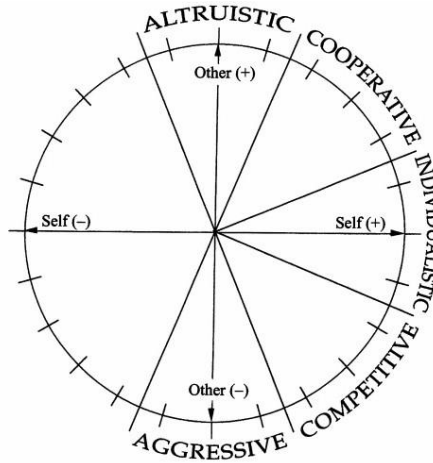
The monetary values of the payoffs are chosen so that, considering a two-dimensional space, the ordered pairs (x, y) are placed at 24 equally spaced points on the circumference of a circle, which is centered at the origin $(0, 0)$ and has a radius of 15 monetary units. This entails that $x^2 + y^2 = 15^2$, where $x + y$ is not constant.

In each scenario the player is supposed to choose between two adjacent points on this circumference and, according to the weight assigned by the subject to the own payoff and to the partner’s payoff, I build a motivational vector indicating the individual’s type.⁵

As in Figure 3.1, individuals are traditionally classified as:

⁵For further details I refer the reader to Offerman et al. (1996); Brosig (2002); Bala-foutas et al. (2013).

Figure 3.1: The value orientation circle by Offerman et al. (1996)



- *aggressive*, with an observed vector lying between -112.5° and -67.5° (measured from the abscissa);
- *competitive*, between -67.5° and -22.5° ;
- *individualistic*, between -22.5° and 22.5° ;
- *cooperative*, between 22.5° and 67.5° ;
- *altruistic*,⁶ between 67.5° and 112.5° .

Moreover, the length of the motivational vector is usually employed as indicator of consistency, that is, it shows how often the subject picked the own-other payoff combination which is the closest to her own motivational vector. If the individual chooses consistently throughout the 24 scenarios, the ensuing vector length is equal to 30. Random choices produce a vector of zero length. In my case, I decide not to exclude observations from the

⁶In Section 3.4 I code the observed angle as *Altruism*.

analysis, since only one subject's vector length is below 7.5, the benchmark used by Brosig (2002).

After completing the Ring-test, the participants are asked to fill in two questionnaires measuring the degree of subjective interdependency perceived during the public goods game and context-independent empathy, respectively.⁷ I place them at this stage of the session for a couple of reasons. On the one hand, the questionnaires are situated after the Ring-test in order to avoid experimenter demand effect. On the other hand, they are filled in before I give feedback about the final payment, thereby holding the attention of the subjects.

Finally, at the end of the session each subject is informed of her own wealth accumulated during the experiment and of the profit relevant to the final payment, which comes from a fifty-fifty lottery between the Ring-test outcome and the period randomly selected from the public goods game.

3.3.3 Hypotheses

In order to check if unexpected low quality of the public good has detrimental effects on the successive decision to contribute, I develop the following hypothesis at the individual level:

H1. Contributions are negatively affected by previous low quality of the public good.

⁷As in Sonnemans et al. (1998), subjective interdependency is measured by means of a seven-point scale for the five following questions: (1) how much influence do you think you had on your own payoffs; (2) how much influence do you think you had on the earnings of other group members; (3) how much influence do you think you had on the decisions of others; (4) how much influence do you think the others had on your earnings; and (5) to what degree do you agree with the statement "In this experiment I and the other group members depend on each other for good results". In Section 3.4 I average out the scores from the five answers under the variable *interdependency*.

The second questionnaire aims at measuring empathy as dispositional trait consisting of both affective and cognitive components. In order to do so, I rely on the Empathy Quotient (EQ), a questionnaire containing 40 empathy items and 20 filler items. Participants can earn 2, 1, or 0 points on each empathy item, so that one's final score ranges from 80 to 0. For further details, see Baron-Cohen and Wheelwright (2004); Lawrence et al. (2004). In Section 3.4 I code the individual score as *Empathy Quotient*.

H1 is assumed to be true especially when the decision makers are allowed to communicate and aggregate information. Indeed, given that the private signals are correlated with the occurrence of high-valued public goods (70%), the sharing of such signals is conducive to high expectations which in case of poor quality would be disappointed.

In addition to H1, I also test some regularities related to efficiency of communicating in threshold public goods games, starting from the outcomes observed by Palfrey et al. (2017) which could be extended to my design. Therefore, I formulate the following hypotheses at the group level:

H2. The likelihood of public good provision is greater in the chat treatment than in the baseline.

H3. The average group earnings are higher in the chat treatment than in the baseline.

H4. The incidence of just one or two group members contributing is lower in the chat treatment than in the baseline.

H5. The incidence of all group members contributing is lower in the chat treatment than in the baseline.

Basically, H2 embodies the traditional positive effect of allowing communication. If supported, this hypothesis should lead to H3, since the lottery associated with the public good provision yields the high value on average with 70% probability.

H4 and H5 reflect the fact that communication improves coordination, that theoretically should occur either on three or on zero contributions. Indeed, by allowing communication, the players' beliefs about the value g of the public good should drive the outcome to either one of the two combinations, since the agreement on four contributions is not credible due to the free-rider problem. Moreover, as I will discuss in the next section, agreements on three contributions are feasible thanks to the presence of a valid coordination device, that is, the rotation of the non-contributor. Thus, I assume that the cases of undercontribution and overcontribution, which are great sources of inefficiency, are reduced in the chat treatment.

3.4 Results

In this section I introduce the results of the experiment by splitting them into three blocks.

First of all, I carry out an overall analysis that checks the performance of H1 and produces preliminary insights into H2. Then I test the hypotheses at the group level. The final subsection analyzes the chat messages and proposes possible explanations.

3.4.1 Overall analysis

Before looking at the overall results, I define the following variables:

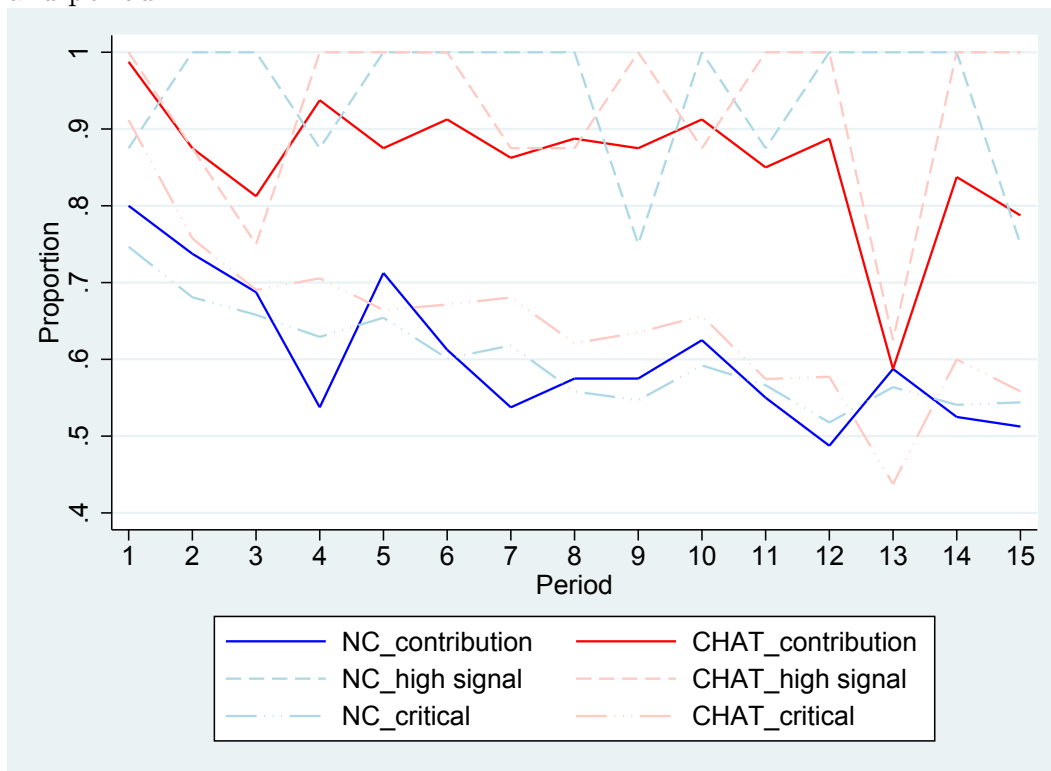
- *contribution*, a dummy equal to 1 if player i decides to contribute, and 0 otherwise.
- *high signal*, a dummy equal to 1 if player i receives a private signal stating that the sample proportion of white balls is high, and 0 otherwise.
- *critical*, the subjective probability of being critical for the threshold achievement as revealed by player i .

Figure 3.2 shows how these variables change over time by comparing the no communication (NC) condition and the chat treatment.

As reported in Table 3.3, the participants in the chat sessions decided to contribute in 85.92% of cases, which according to a Chi-squared test is significantly different from 60.42% displayed by players who were not allowed to interact ($p < 0.01$). This discrepancy is clear-cut over periods and cannot be explained by a hypothetical difference in randomness of high signals between treatments. Indeed, high signals were generated with similar elevated frequencies in the two treatments, as the same nonparametric test confirms. Moreover, the subjective probability of being critical is significantly lower in the baseline⁸ ($p < 0.01$) and seems to be positively correlated with the contribution choice, as already observed in Offerman et al. (1996). Indeed,

⁸In this case, a one-tailed Wilcoxon rank sum test is used.

Figure 3.2: Contributions, high signals and subjective pivotality by treatment and period



both the two variables tend to decrease over time, with this process being especially pronounced in *CHAT_critical*. This is arguably due to the fact that, as I will show in subsection 3.4.3, the group members often clinched deals on four contributors, thus feeling less critical for the threshold achievement even in case of contribution.

An additional insight of Figure 3.2 is that the subjects relied heavily on the private signals in order to make the contribution decision. This is particularly the case of chat treatment, where the negative peaks at period 13 are emblematic: few high signals entail a limited number of contributions.

Such a preliminary remark is confirmed by the random effects probit regressions in Table 3.4, where I tackle potential dependence across players by clustering standard errors at the group level. Indeed, column (1) shows that the coefficient of *high signal* is positive and significant at the 1% level. In

Table 3.3: Frequencies over periods

	No communication	Chat
<u>Periods 1-5</u>		
Contribution	69.50	89.75***
High signal	95	92.50
Critical	67.38	74.59***
Observations	400	400
<u>Periods 6-10</u>		
Contribution	58.50	89***
High signal	95	92.50
Critical	58.33	65.30***
Observations	400	400
<u>Periods 11-15</u>		
Contribution	53.25	79***
High signal	92.50	92.50
Critical	54.66	54.95
Observations	400	400
<u>All periods</u>		
Contribution	60.42	85.92***
High signal	94.17	92.50
Critical	60.12	64.94***
Observations	1,200	1,200

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

other words, the subjects closely followed the suggestions given by their own signals. But in particular, the positive and significant coefficient ($p < 0.01$) of the treatment dummy *chat* highlights the crucial role played by pre-play communication. This effect is genuine, because in the regression I control for information aggregation by including a variable for the number of signals one had access to. This result at the individual level already hints at the validity of H2. At the same time, the contribution decision is also driven by subjective pivotality and the measure of altruism obtained through the Ring-test ($p < 0.01$ in both cases).

At this point, I test H1 by checking the role of the group ability to reach the threshold in the previous period but, quite surprisingly, the coefficients of *high-valued PG* and *low-valued PG* are not statistically significant. This implies that the successful provision of the public good, whether its quality is good or poor, has no consequences for the contribution choice in the successive period, thereby leading to rejection of H1:

Result 1. *Contributions are not negatively affected by previous low quality of the public good (reject H1).*

Conjecturing that the low value of the public good might be more salient in the chat treatment where information aggregation raises expectations about quality, in column (2) I add the interaction terms but still no effects are found. Column (3) just leaves out the lagged variables and the other predictors which are not significant, thus recovering observations. The results remain unchanged.

All in all, given that the regressor *altruism* has an impact on the contribution decision, I conclude this preliminary overview by wondering whether the discrepancy in the proportion of contributions can be ascribed to greater presence of cooperative and altruistic players in the chat treatment. Therefore, in Table 3.5 I examine participants' social orientation as revealed by the Ring-test, but a Fisher's exact test is not able to reject the hypothesis of no differences in the distribution in categories ($p = 0.443$).

Table 3.4: Explaining contributions

Dependent variable: contribution decision			
	(1)	(2)	(3)
Chat	0.843*** (0.315)	1.077*** (0.405)	0.905*** (0.323)
<i>High-valued PG</i> (t - 1)	0.030 (0.125)	0.058 (0.159)	
<i>Low-valued PG</i> (t - 1)	-0.097 (0.121)	-0.013 (0.150)	
<i>High-valued PG * Chat</i> (t - 1)		-0.232 (0.230)	
<i>Low-valued PG * Chat</i> (t - 1)		-0.369 (0.245)	
Trend	-0.027*** (0.010)	-0.029*** (0.010)	-0.034*** (0.010)
Signals accessed	0.213** (0.088)	0.216** (0.088)	0.209** (0.089)
High signal	1.107*** (0.180)	1.082*** (0.180)	1.166*** (0.173)
Critical	0.022*** (0.004)	0.022*** (0.004)	0.022*** (0.003)
Female	-0.419** (0.202)	-0.417** (0.202)	-0.376* (0.197)
Risk aversion	0.024 (0.063)	0.023 (0.063)	
Altruism	0.009** (0.003)	0.009** (0.003)	0.008*** (0.003)
Empathy Quotient	-0.001 (0.011)	-0.000 (0.011)	
Interdependency	0.190 (0.133)	0.190 (0.133)	
<i>Constant</i>	-2.798*** (0.800)	-2.806*** (0.790)	-1.976*** (0.345)
Log pseudolikelihood	-851.388	-850.467	-891.310
Wald χ^2	288.750	289.510	316.280
Prob > χ^2	0.000	0.000	0.000
Observations	2,240	2,240	2,400

Coefficient estimates from random effects probit models, with standard errors in parentheses clustered at the group level. *Chat* is the treatment dummy. *Trend* is a linear time trend that starts from 0. *High-valued PG* (*t* - 1) is a dummy equal to 1 if subject's group reached the threshold and obtained a high-valued public good in the previous period. *Low-valued PG* (*t* - 1) is a dummy equal to 1 if subject's group reached the threshold and obtained a low-valued public good in the previous period. The omitted category is the case in which subject's group did not reach the threshold in the previous period.

***Significant at the 1% level. 71

**Significant at the 5% level.

*Significant at the 10% level.

Table 3.5: Social orientation

	No communication	Chat
Aggressive	0	1
Competitive	3	6
Individualistic	41	44
Cooperative	35	29
Altruistic	1	0
Observations	80	80

Absolute frequencies.

3.4.2 Group outcomes

Now I move on to the group level analysis and test the hypotheses related to efficiency.

Considering each group as independent observation, in Table 3.6 I compare the frequency of public good provision and the average group earnings in the two treatments by means of a one-tailed Wilcoxon rank sum test.

As the considerable discrepancy in contributions at the individual level already suggested, the overall provision rate is significantly greater ($p < 0.01$) when the subjects are free to exchange messages. The outcome is not different if I take into account any subset of periods. Indeed, in the chat treatment the decline of public good provision is really subtle and the groups reached the threshold on average in 92% of cases, with respect to 56% in the baseline. On the contrary, I notice a pronounced decrease of provision rate in the absence of communication.

Therefore, I can conclude the following:

Result 2. *The likelihood of public good provision is significantly greater with communication than without communication (support for H2).*

Given this result, it is not surprising to find out that also the average group earnings are significantly higher at the 1% level in the chat treatment.⁹

⁹The average group earnings in Table 3.6 include two of the three profit components

Indeed, every time that the public good is provided, the associated lottery produces the high value on average with 70% probability. In each period, the groups in the communication sessions gained on average 253.58 ECU versus 179.15 ECU of the groups that were not allowed to chat. The mean difference between the two treatments is not that evident at the beginning of the game but goes beyond 110 ECU in the final periods, because of the provision rate downfall in the no communication sessions.

Hence, H3 is supported as well:

Result 3. *The average group earnings are significantly higher in the chat sessions than in the no communication sessions (support for H3).*

So far so good: the communication effect has given evidence of being beneficial at the group level, in accordance with the literature.

At this point, I wonder if pre-play communication succeeded in lowering the sources of inefficiency that I detected in subsection 3.3.1. The answer is contained in Table 3.7, which displays the frequency distributions of the number of contributors in the two treatments. Taking all periods into consideration, the incidence of just one or two group members contributing (wasteful undercontribution) is significantly lower at the 1% level, when communication is allowed. Indeed, this occurred only in 5% of cases with respect to 34% in the baseline, with the significant difference being consistent over periods. Therefore, H4 is supported.

The striking result is that, when the opportunity to make unbinding deals was given, the subjects massively preferred to coordinate on all group members contributing, and these agreements were often successful despite the free-rider problem. Indeed, not only H5 is not supported, but also its opposite is true. The incidence of all group members contributing (wasteful overcontribution) turns out to be significantly higher at the 1% level in the chat treatment (59%) than in the baseline (18%), with this outcome being visible in any subset of periods and at odds with the findings in Palfrey et al. (2017). As the reader will see in the next subsection thanks to chat analysis,

of the game, that is, the initial endowment and the net return (the gain coming from the incentivized question here is not considered).

Table 3.6: Frequency of public good provision and average group earnings

	No communication	Chat
<u>Periods 1-5</u>		
Provision rate	0.68	0.96***
<i>High value</i>	0.39	0.58
<i>Low value</i>	0.29	0.38
Group earnings	190.70	230.35**
Observations	20	20
<u>Periods 6-10</u>		
Provision rate	0.52	0.95***
<i>High value</i>	0.33	0.66
<i>Low value</i>	0.19	0.29
Group earnings	178.90	249.80***
Observations	20	20
<u>Periods 11-15</u>		
Provision rate	0.47	0.86***
<i>High value</i>	0.27	0.76
<i>Low value</i>	0.20	0.10
Group earnings	167.85	280.60***
Observations	20	20
<u>All periods</u>		
Provision rate	0.56	0.92***
<i>High value</i>	0.33	0.67
<i>Low value</i>	0.23	0.25
Group earnings	179.15	253.58***
Observations	20	20

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

the percentage of successful agreements on four contributors is surprisingly high as well.

Thus, I draw the following conclusions:

Result 4. *The incidence of just one or two group members contributing is significantly lower in the chat treatment than in the baseline (support for H_4).*

Result 5. *The incidence of all group members contributing is not significantly lower in the chat treatment than in the baseline (reject H_5). On the contrary, it is significantly higher. The four-contributor category becomes the modal class when communication is allowed.*

Finally, the incidence of groups without contributions is similar in the two treatments, as confirmed by the Wilcoxon rank sum test. Here I want to emphasize that in the baseline, clearly, coordination on zero contributions can emerge only by chance, but in the chat treatment it represents an efficient option, in case the group members anticipate a low value by means of the signals. Therefore, the minimal percentages throughout the fifteen periods might be a sign of high expectations about the quality of the public good.

To sum up, in my setting the opportunity to communicate significantly increases public good provision by reducing inefficiency that comes from wasteful undercontribution. The observed higher provision rate is fostered by the fact that the group members hardly ever decided to coordinate on zero contributions when they were free to communicate. Moreover, the players in this condition seemed to disregard the free-rider issue and often ended up overcontributing. This inefficient choice may be more pervasive than observed, in case the outcomes with three group members contributing are due to defectors and not to stated agreements.

3.4.3 Chat analysis and further insights

At this point, I analyze the chat messages in order to reach a better understanding of the group outcomes.

Table 3.7: Frequency distributions of number of contributors

Contributors	No communication	Chat
<u>Periods 1-5</u>		
0	0.04	0.03
1 or 2	0.28	0.01***
3	0.45	0.27**
4	0.23	0.69***
Efficient outcomes (0 or 3)	0.49	0.30**
Observations	20	20
<u>Periods 6-10</u>		
0	0.10	0.01
1 or 2	0.38	0.04***
3	0.36	0.31
4	0.16	0.64***
Efficient outcomes (0 or 3)	0.46	0.32*
Observations	20	20
<u>Periods 11-15</u>		
0	0.16	0.05
1 or 2	0.37	0.09***
3	0.33	0.42
4	0.14	0.44***
Efficient outcomes (0 or 3)	0.49	0.47
Observations	20	20
<u>All periods</u>		
0	0.10	0.03
1 or 2	0.34	0.05***
3	0.38	0.33
4	0.18	0.59***
Efficient outcomes (0 or 3)	0.48	0.36*
Observations	20	20

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

Since the design is characterized by partner matching, the players were strongly encouraged to speak out and clinch long-lasting deals. I notice a not negligible tendency to come up with a group strategy in the early rounds and subsequently carry it on with tacit agreement. This prevents me from interpreting individual silence as lack of a common action profile, so that I make the realistic assumption that the 20 groups in the communication sessions reached agreement all the time.

Table 3.8: Agreements on the action profiles

Periods	Agreed number of contributors			
	0	1 or 2	3	4
1-5	3 (3)	0 (0)	13 (13)	69 (84)
6-10	1 (1)	0 (0)	12 (13)	63 (86)
11-15	5 (7)	0 (0)	14 (17)	43 (76)
All periods	9 (11)	0 (0)	39 (43)	175 (246)

Number of times the group members succeeded in carrying out the agreed action profile in the chat treatment (number of agreements in parentheses).

Table 3.8 suggests that the problem of overcontribution is greater than what could be observed from the mere number of group contributors. Indeed, 82% of the agreements (246 out of 300 interactions) were reached on all group members contributing, even if actual overcontribution amounted only to 59% of cases. Despite the non-credible nature of these agreements and the concrete threat of free-riding, the percentage of successful deals on four contributors is surprisingly high, and equal to 71% (175 out of 246). This is consistent with the fact that partner matching tends to reduce the occurrence of free-riding (Solow and Kirkwood, 2002). At the same time, such an outcome might be driven by the pursuit of symmetric payoffs within the group, a condition which is generally conducive to coordination (Crawford et al., 2008).

Anyway, I observe that the proportion of groups with three players contributing (33%) is moderately overstated by the free-rider issue, considering that only 14% of the deals (43 out of 300) were made on three contributors. It goes without saying that the agreements on this action profile were suc-

cessfully carried out very often, up to 91% of cases (39 out of 43). Scarce convergence is detected on zero contributions, whereas none on the remaining categories.

Now I run a Wilcoxon rank sum test on the variable *interdependency*, which is a proxy for group identity (Jackson and Smith, 1999) coded as in subsection 3.3.2, and I discover that players perceived to be significantly more interconnected with the other group members in the chat treatment than in the baseline ($z = -2.832$, $p < 0.01$). This result can be explained through the common fate hypothesis, according to which “when the members of a group are exposed to the same risk, group identification is facilitated” (Corazzini and Sugden, 2011). Moreover, greater group identity is expected to make team reasoning more salient, whose related theories claim that when a player identifies with the group, she will detect and implement a strategy that achieves the maximum group payoff, thereby playing her own part within such a joint action (Bacharach, 1999; Sugden, 2000, 2003). In my experimental design, the partner matching allows to maximize the group earnings in every period by means of a rotation of the only one member not contributing that would solve the problem of free-riding as well. Therefore, in case the players foresee high quality of the public good and accordingly agree on reaching the threshold, theories of team reasoning would predict the implementation of such a rotation.

Table 3.9: The actual strategies implemented

Strategy	
Signal-based	128 (154)
Rotation-based	27 (28)
<i>Others</i>	68 (118)
Total	223 (300)

Number of times the group members succeeded in carrying out the agreed strategy in the chat treatment (number of strategies in parentheses).

Actually, Table 3.9 shows that just 9% of the actual strategies (28 out of 300) were grounded in the rotation of the non-contributor, with a success

rate equal to 96% (27 out of 28). On the one hand, the scarce use of this coordination device is somewhat unexpected. On the other hand, its high effectiveness is fully predictable because this strategy involves exclusively agreements on three contributors.

In 51% of cases (154 out of 300) the players came up with a strategy based on the systematic sharing of all four private signals of the group, this action profile aiming to reduce uncertainty. Furthermore, in 97% of these interactions (150 out of 154) all the group members honestly reported the true values and 83% of the signal-based strategies (128 out of 154) were successfully carried out.

The residual class contains 39% of the observations (118 out of 300) and is mainly constituted by interactions where only some of the group members disclosed private information. Therefore, the lower success rate (58%) may be justified by the lack of a precise strategy.

All in all, the groups did not maximize earnings but preferred to pursue symmetric payoffs and minimize ambiguity through the sharing of signals period by period, a strategy which is inefficient because not compatible with the rotation of the non-contributor.

3.5 Conclusions

In this chapter I offered experimental evidence on the impact of communication on the provision of public goods whose value is ambiguous. I drew inspiration from the case of joint research projects where the quality of the final output cannot be inferred in advance. Within a context of uncertainty, communication could lead to inefficient outcomes and, therefore, the current chapter aimed to analyze such dynamics in a controlled environment.

The analysis at the group level emphasized that, in accordance with related literature, communication significantly increased public good provision by reducing inefficiency that comes from wasteful undercontribution. The observed higher provision rate in the chat treatment was fostered by the fact that the group members hardly ever decided to coordinate on zero contributions when they were free to communicate. Indeed, contrary to my hypothesis based on Bell (1985)'s theory of disappointment in decision mak-

ing under uncertainty, the subjects were not let down when they ran into low quality of the public good, after reaching the threshold. Moreover, the players in this experimental condition seemed to neglect the free-rider issue and often ended up overcontributing, in contrast with the study by Palfrey et al. (2017).

After chat analysis, I discovered that the issue of overcontribution was even more pervasive than what could be observed from the mere number of group contributors, due to moderate presence of defectors within agreements on all members contributing. Nevertheless, the deals on this action profile were often successful, arguably thanks to subjects' willingness to pursue symmetric payoffs as well as to pronounced group identity generated by the partner matching and the common fate hypothesis. Despite the presence of a valid coordination device as the rotation of the non-contributor, the group members privileged agreements on four contributions and the implementation of strategies based on the systematic disclosure of all four private signals. Since the players preferred to minimize uncertainty than to maximize the group earnings, this chapter makes an empirical contribution to the literature on team reasoning by highlighting cases where the related theories cannot be supported. As a result, in the presence of uncertainty satisficing is more salient than optimizing, this insight paving the way for promising investigations.

To conclude, regarding possible concrete applications I stress the fact that the experiment was not framed, so that a number of compatible environments can benefit from the results. Going back to the introductory example, the findings of the chapter may explain the tendency to observe research projects including a large number of scholars through the willingness not to be outperformed. Such a strategy could be undertaken in order to reduce uncertainty about the final output as well. Instead, it is harder to say how serious the inefficiency from overcontribution is, the answer arguably depending on the various real-world situations. Definitely, further research is needed to reinforce this view.

Appendix 3A. Experimental instructions

[The following instructions are translated from Spanish. They were provided on paper and were the same in both treatments, except for the paragraph in bold related to the chat and for the value of the show-up fee.]

TASK 1

Thank you very much for being here. All participants have the same instructions: please read them carefully and, if you have questions or doubts, raise your hand and a member of the staff will answer you individually. During the session, it is strictly forbidden to speak with the other participants.

The experimental currency unit is the ECU, with an exchange rate of 5 ECU = 1 €. The current task consists in making a number of decisions during fifteen periods. In this experiment it is not possible to make losses, your final profit in euro depending on your own decision made in a period randomly chosen by the PC.

In each period you will interact with other three participants in this room. The four of you form a group that will remain the same in all periods. The identity of the other group members will not be revealed to you throughout the session.

Every period, you have to decide whether you want to contribute to a public good.

If you decide to contribute, your amount of the contribution will be fixed and equal to your entire endowment, that is, 50 ECU (cost of contributing). The share of public good that you are going to win will depend on whether the other group members decide to contribute as well. If overall less than three group members contribute, nobody within the group will win anything and you will just lose your contribution.

Nevertheless, if overall at least three group members contribute, the share of public good that you are going to win will depend on a random draw. There is a box containing 100 balls, where each ball has the same probability of being white (this holds throughout the game), but no player knows this probability. A ball will be drawn from the box: if the ball is white, the share of public good that will be assigned to you is equal to 80 ECU and, therefore, your net return will be equal to 30 ECU. If the ball is not white, the share of public good that will be assigned to you is equal to 20 ECU and your net return will be equal to -30 ECU.

Before making her own contribution choice, each group member has the opportunity to individually draw a sample of 10 balls from the box, thereby receiving a message which can be of two kinds:

- The sample proportion of white balls is *High* (greater than or equal to 5/10).
- The sample proportion of white balls is *Low* (less than 5/10).

If you decide not to contribute, two scenarios are possible:

- Another group member decides not to contribute, and in this case neither you nor the other group members will win anything. You will sustain a cost of 15 ECU, but the rest of your initial endowment will be safe.
- You are the only group member who decides not to contribute, and in this case your share of public good will depend on the random draw, as already explained before. The difference is that your cost, in this case, will be equal to 15 ECU.

Moreover, in each period you will be asked an incentivized question about your subjective probability (from 0 to 100%) that exactly other two group members contribute to the public good in that period. You can win either 5 extra ECU or nothing according to the ex post accuracy of your prediction. This part of your profit will be calculated by means of a Binary Lottery Procedure:

- Being θ your subjective probability that exactly two out of the other three group members contribute, the probability of winning 5 extra ECU is equal to $P(\theta | C) = 100 - 100(1 - \theta/100)^2$ in case actually two out of the other three group members contribute. Otherwise, that probability is equal to $P(\theta | NC) = 100 - 100(0 - \theta/100)^2$.

For instance, if your subjective probability that exactly two out of the other three group members contribute is equal to 80%, and this actually occurs, you win 5 extra ECU with the following probability: $P(80 | C) = 100 - 100(1 - 80/100)^2 = 100 - 4 = 96\%$.

[The following paragraph in bold is present only in the chat treatment.]

Before making the contribution decision, you have the opportunity to chat with the other group members for one minute and a half, precisely. During this interval, you are allowed exchange messages. Please take into account the following points:

- **All the messages sent by each player will reach all the group members.**
- **It is forbidden to write messages whose aim is to identify the other group members.**
- **It is forbidden to write messages whose aim is to reveal your own identity to the rest of the group.**
- **Your messages must be related to your decision-making process in this experiment.**

Once time expires, the chat will close.

In every period, your profits will be the sum of three parts:

1. Your initial endowment of 50 ECU.

2. Your net return, that is, your share of public good minus the cost of (not) contributing, this net return ranging from -50 to 65 ECU.
3. The earnings from the incentivized question: either 5 ECU or nothing.

Remember: only one period and its corresponding profit will be relevant to the final payment in euro.

Thank you for participating!

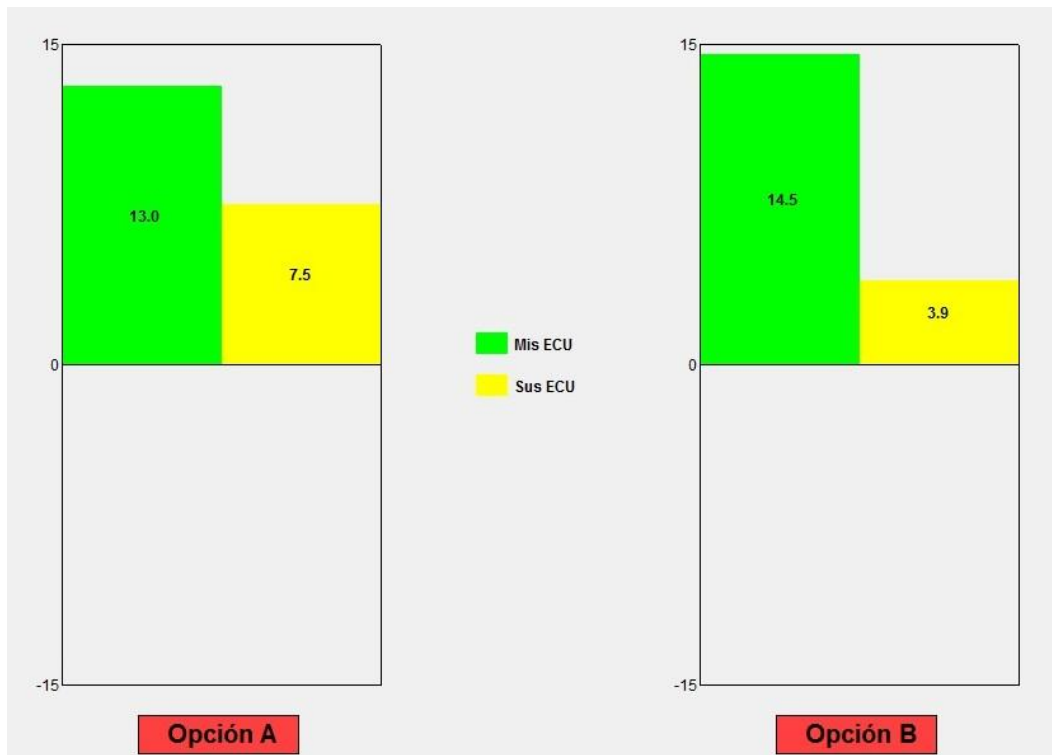
TASK 2

Today's experiment also consists of another task which is unrelated to that one just carried out. At the end of the session you will be paid via either one of the two, with the relevant task being chosen randomly by the PC, in addition to a fixed amount of 5 euro that you cannot lose in any case. *[The value of the show-up fee was increased to 7 euro in the chat treatment.]*

In this second task your initial endowment is equal to 40 ECU and you are going to be randomly matched with another participant of the session. Neither of the two will know the partner's identity.

You are given 24 scenarios and, in each of them, you are supposed to choose either one of two options: "A" and "B". Each option assigns either a positive or a negative quantity of ECU to you and your partner. The amount of ECU that you are going to win in each scenario will depend on the options chosen by you and your partner, the structure of the 24 scenarios being the same. Throughout this task, you will not be receiving any information about the option picked by your partner: you will just know about the overall amount obtained at the end, after having made the 24 choices.

Below we provide you with an example of the decision-making process just described. Suppose that you end up with a situation like the following one:



There are four possible results:

1. If you choose option "A" and your partner picks option "B" (something that you are not able to know), you will obtain 13 ECU (since you have chosen option "A") plus 3.9 ECU (since your partner has chosen option "B"), this scenario thereby generating a total amount of 16.9 ECU for you and 22 ECU for your partner.
2. If you both choose option "A", you will obtain 13 ECU (since you have chosen option "A") plus 7.5 ECU (since your partner has chosen option "A"), this scenario thereby generating a total amount of 20.5 ECU each.
3. If you choose option "B" and your partner picks option "A", you will obtain 14.5 ECU (since you have chosen option "B") plus 7.5 ECU (since your partner has chosen option "A"), this scenario thereby generating a total amount of 22 ECU for you and 16.9 ECU for your partner.
4. If you both choose option "B", you will obtain 14.5 ECU (since you have chosen option "B") plus 3.9 ECU (since your partner has chosen option "B"), this scenario thereby generating a total amount of 18.4 ECU each.

The PC will automatically calculate the amounts obtained in each scenario, but you will not be given information about the results until you make all 24 decisions.

In case this task is that one randomly chosen in order to determine your profits at the end of the session, then you will be paid the sum of your earnings in the 24 scenarios, by means of the exchange rate $5 \text{ ECU} = 1 \text{ €}$.

In a while the game is going to start: please click the button related to your preferred option in order to make a choice in the current scenario.

Bibliography

- Abdellaoui, M., Driouchi, A., and l'Haridon, O. (2011). Risk aversion elicitation: reconciling tractability and bias minimization. *Theory and Decision*, 71(1):63–80.
- Agastya, M., Menezes, F., and Sengupta, K. (2007). Cheap talk, efficiency and egalitarian cost sharing in joint projects. *Games and Economic Behavior*, 60(1):1–19.
- Alberti, F. and Cartwright, E. J. (2016). Full agreement and the provision of threshold public goods. *Public Choice*, 166(1):205–233.
- Andrade, E. B. and Ariely, D. (2009). The enduring impact of transient emotions on decision making. *Organizational Behavior and Human Decision Processes*, 109(1):1–8.
- Bacharach, M. (1999). Interactive team reasoning: A contribution to the theory of co-operation. *Research in Economics*, 53(2):117–147.
- Bailey, R. W., Eichberger, J., and Kelsey, D. (2005). Ambiguity and public good provision in large societies. *Journal of Public Economic Theory*, 7(5):741–759.
- Balafoutas, L., Kocher, M. G., Putterman, L., and Sutter, M. (2013). Equality, equity and incentives: An experiment. *European Economic Review*, 60:32–51.
- Baron-Cohen, S. and Wheelwright, S. (2004). The empathy quotient: an investigation of adults with asperger syndrome or high functioning aut-

- ism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34(2):163–175.
- Barrett, S. and Dannenberg, A. (2012). Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences*, 109(43):17372–17376.
- Baumeister, R. E., Bratslavsky, E., Muraven, M., and Tice, D. M. (1998). Ego Depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74(5):1252–1265.
- Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research*, 33(1):1–27.
- Bicchieri, C. and Lev-On, A. (2007). Computer-mediated communication and cooperation in social dilemmas: an experimental analysis. *Politics, Philosophy & Economics*, 6(2):139–168.
- Binswanger, H. P. (1980). Attitudes toward risk: Experimental measurement in rural india. *American Journal of Agricultural Economics*, 62(3):395–407.
- Blais, A.-R. and Weber, E. U. (2006). A domain-specific risk-taking (dospert) scale for adult populations. *Judgment and Decision Making*, 1(1):33–47.
- Blanco, M., Engelmann, D., Koch, A. K., and Normann, H.-T. (2010). Belief elicitation in experiments: is there a hedging problem? *Experimental Economics*, 13(4):412–438.
- Bochet, O., Page, T., and Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization*, 60(1):11–26.
- Brosig, J. (2002). Identifying cooperative behavior: some experimental results in a prisoner’s dilemma game. *Journal of Economic Behavior & Organization*, 47(3):275–290.
- Calford, E. (2017). Uncertainty aversion in game theory: Experimental evidence. *Purdue University, Department of Economics, Working Paper No. 1291*.

- Camerer, C. F. and Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19(1-3):7–42.
- Campos-Vazquez, R. M. and Cuijty, E. (2014). The role of emotions on risk aversion: a Prospect Theory experiment. *Journal of Behavioral and Experimental Economics*, 50:1–9.
- Caporael, L. R., Dawes, R. M., Orbell, J. M., and Van De Kragt, A. (1989). Selfishness examined: Cooperation in the absence of egoistic incentives. *Behavioral and Brain Sciences*, 12(4):683–699.
- Capraro, V. et al. (2018). Gender differences in lying in sender-receiver games: A meta-analysis. *Judgment and Decision Making*, 13(4):345–355.
- Charness, G., Gneezy, U., and Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization*, 87:43–51.
- Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, 14(1):47–83.
- Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, 19(22):3127–3131.
- Chou, K.-L., Lee, T., and Ho, A. H. (2007). Does mood state change risk taking tendency in older adults? *Psychology and Aging*, 22(2):310–318.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: L. Erlbaum Associates.
- Cohn, A., Engelmann, J., Fehr, E., and Maréchal, M. A. (2015). Evidence for countercyclical risk aversion: An experiment with financial professionals. *American Economic Review*, 105(2):860–885.
- Colasante, A., Marini, M. M., and Russo, A. (2017). Incidental emotions and risk-taking: An experimental analysis. *MPRA Paper 82767, University Library of Munich, Germany*.

- Conte, A., Levati, M. V., and Nardi, C. (2018). Risk preferences and the role of emotions. *Economica*, 85(338):305–328.
- Cooper, R., DeJong, D. V., Forsythe, R., and Ross, T. W. (1992). Communication in coordination games. *The Quarterly Journal of Economics*, 107(2):739–771.
- Corazzini, L. and Sugden, R. (2011). Common fate, game harmony and contributions to public goods: experimental evidence. *International Review of Economics*, 58(1):43–52.
- Cox, C. A. (2015). Cursed beliefs with common-value public goods. *Journal of Public Economics*, 121:52–65.
- Crawford, V. P., Gneezy, U., and Rottenstreich, Y. (2008). The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures. *American Economic Review*, 98(4):1443–58.
- Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431–1451.
- Crosetto, P. and Filippin, A. (2016). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics*, 19(3):613–641.
- Cubitt, R. P., Starmer, C., and Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, 1(2):115–131.
- Dawes, R. M., Orbell, J. M., Simmons, R. T., and Van De Kragt, A. (1986). Organizing groups for collective action. *American Political Science Review*, 80(4):1171–1185.
- Dhami, S. (2016). *The foundations of behavioral economic analysis*. Oxford University Press.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

- Drichoutis, A. C. and Nayga Jr, R. M. (2013). Eliciting risk and time preferences under induced mood states. *The Journal of Socio-Economics*, 45:18–27.
- Eichberger, J. and Kelsey, D. (2002). Strategic complements, substitutes, and ambiguity: the implications for public goods. *Journal of Economic Theory*, 106(2):436–466.
- Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4):643–669.
- Elster, J. (1998). Emotions and economic theory. *Journal of Economic Literature*, 36(1):47–74.
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, 14(4):583–610.
- Fehr-Duda, H., Epper, T., Bruhin, A., and Schubert, R. (2011). Risk and rationality: The effects of mood and decision rules on probability weighting. *Journal of Economic Behavior & Organization*, 78(1-2):14–24.
- Fessler, D. M., Pillsworth, E. G., and Flamson, T. J. (2004). Angry men and disgusted women: An evolutionary approach to the influence of emotions on risk taking. *Organizational Behavior and Human Decision Processes*, 95(1):107–123.
- Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.
- Fischer, P., Greitemeyer, T., and Frey, D. (2007). Ego depletion and positive illusions: Does the construction of positivity require regulatory resources? *Personality and Social Psychology Bulletin*, 33(9):1306–1321.
- Forgas, J. P. (1995). Mood and Judgment: the Affect Infusion Model (AIM). *Psychological Bulletin*, 117(1):39–66.
- Friedman, D. and Sunder, S. (1994). *Experimental methods: A primer for economists*. Cambridge University Press.

- George, J. M. and Dane, E. (2016). Affect, emotion, and decision making. *Organizational Behavior and Human Decision Processes*, 136:47–55.
- Grable, J. E. and Roszkowski, M. J. (2008). The influence of mood on the willingness to take financial risks. *Journal of Risk Research*, 11(7):905–923.
- Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125.
- Guiso, L. and Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association*, 6(6):1109–1150.
- Guiso, L., Sapienza, P., and Zingales, L. (2018). Time varying risk aversion. *Journal of Financial Economics*, 128(3):403–421.
- Güth, W., Levati, M. V., and Soraperra, I. (2015). Common and private signals in public goods games with a point of no return. *Resource and Energy Economics*, 41:164–184.
- Hanoch, Y., Johnson, J. G., and Wilke, A. (2006). Domain specificity in experimental measures and participant recruitment: An application to risk-taking behavior. *Psychological Science*, 17(4):300–304.
- Harrison, G. W., Martínez-Correa, J., and Swarthout, J. T. (2014). Eliciting subjective probabilities with binary lotteries. *Journal of Economic Behavior & Organization*, 101:128–140.
- Hedges, L. and Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.
- Heilman, R. M., Crişan, L. G., Houser, D., Miclea, M., and Miu, A. C. (2010). Emotion regulation and decision making under risk and uncertainty. *Emotion*, 10(2):257–265.
- Hertwig, R. and Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences*, 24(3):383–451.

- Hey, J. D. and Lee, J. (2005). Do subjects separate (or are they sophisticated)? *Experimental Economics*, 8(3):233–265.
- Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655.
- Hsee, C. K. and Weber, E. U. (1999). Cross-national differences in risk preference and lay predictions. *Journal of Behavioral Decision Making*, 12(2):165–179.
- Hu, Y., Wang, D., Pang, K., Xu, G., and Guo, J. (2015). The effect of emotion and time pressure on risk decision-making. *Journal of Risk Research*, 18(5):637–650.
- Isen, A. M. and Patrick, R. (1983). The effect of positive feelings on risk taking: When the chips are down. *Organizational Behavior and Human Performance*, 31(2):194–202.
- Jackson, J. W. and Smith, E. R. (1999). Conceptualizing social identity: A new framework and evidence for the impact of different dimensions. *Personality and Social Psychology Bulletin*, 25(1):120–135.
- Johnson, N. D. and Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5):865–889.
- Kaivanto, K. and Kroll, E. B. (2014). Alternation bias and reduction in St. Petersburg gambles: An experimental investigation. *Lancaster University, Economics Working Paper Series, Vol. 2014 No. 14*.
- Kamstra, M. J., Kramer, L. A., and Levi, M. D. (2003). Winter blues: A sad stock market cycle. *American Economic Review*, 93(1):324–343.
- Keenan, D. C., Kim, I., and Warren, R. S. (2006). The private provision of public goods under uncertainty: A symmetric-equilibrium approach. *Journal of Public Economic Theory*, 8(5):863–873.
- Kelsey, D. and Le Roux, S. (2017). Dragon slaying with ambiguity: theory and experiments. *Journal of Public Economic Theory*, 19(1):178–197.

- Kim, D.-Y. and Park, J. (2010). Cultural differences in risk: The group facilitation effect. *Judgment and Decision Making*, 5(5):380.
- Kim, M. Y. and Kanfer, R. (2009). The joint influence of mood and a cognitively demanding task on risk-taking. *Motivation and Emotion*, 33(4):362–372.
- Kliger, D. and Levy, O. (2003). Mood-induced variation in risk preferences. *Journal of Economic Behavior & Organization*, 52(4):573–584.
- Kuhnen, C. M. and Knutson, B. (2011). The influence of affect on beliefs, preferences, and financial decisions. *Journal of Financial and Quantitative Analysis*, 46(3):605–626.
- Lane, T. (2017). How does happiness relate to economic behaviour? A review of the literature. *Journal of Behavioral and Experimental Economics*, 68:62–78.
- Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., and David, A. S. (2004). Measuring empathy: reliability and validity of the empathy quotient. *Psychological Medicine*, 34(5):911–920.
- Ledyard, J. O. (1995). Public goods: A survey of experimental research. In Kagel, J. and Roth, A., editors, *Handbook of Experimental Economics*. Princeton: Princeton University Press.
- Lee, C. J. and Andrade, E. B. (2015). Fear, excitement, and financial risk-taking. *Cognition and Emotion*, 29(1):178–187.
- Lerner, J. S. and Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, 81(1):146–159.
- Lerner, J. S., Li, Y., Valdesolo, P., and Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, 66:799–823.
- Liebrand, W. B. (1984). The effect of social motives, communication and group size on behaviour in an n-person multi-stage mixed-motive game. *European journal of social psychology*, 14(3):239–264.

- Loewenstein, G., O'Donoghue, T., and Bhatia, S. (2015). Modeling the interplay between affect and deliberation. *Decision*, 2(2):55–81.
- Matell, M. S. and Jacoby, J. (1971). Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity. *Educational and Psychological Measurement*, 31(3):657–674.
- Mayer, J. D., Allen, J. P., and Beauregard, K. (1995). Mood inductions for four specific moods: A procedure employing guided imagery vignettes with music. *Journal of Mental Imagery*, 19:133–150.
- Moffatt, P. G. (2015). *Experimentics: Econometrics for experimental economics*. Macmillan International Higher Education.
- Morris, S. B. and DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7(1):105–125.
- Offerman, T., Schram, A., and Sonnemans, J. (1998). Quantal response models in step-level public good games. *European Journal of Political Economy*, 14(1):89–100.
- Offerman, T., Sonnemans, J., and Schram, A. (1996). Value orientations, expectations and voluntary contributions in public goods. *Economic Journal*, 106:817–845.
- Oosterbeek, H., Sloof, R., and Van De Kuilen, G. (2004). Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics*, 7(2):171–188.
- Ostrom, E., Gardner, R., and Walker, J. (1994). *Rules, games, & common-pool resources*. University of Michigan Press.
- Palfrey, T. and Rosenthal, H. (1991). Testing for effects of cheap talk in a public goods game with private information. *Games and Economic Behavior*, 3(2):183–220.

- Palfrey, T., Rosenthal, H., and Roy, N. (2017). How cheap talk enhances efficiency in threshold public goods games. *Games and Economic Behavior*, 101:234–259.
- Raghunathan, R. and Pham, M. T. (1999). All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes*, 79(1):56–77.
- Rick, S. and Loewenstein, G. (2008). The role of emotion in economic behavior. In Lewis, M., Haviland-Jones, J. M., and Feldman Barrett, L., editors, *Handbook of Emotions, Third Edition*, pages 138–158. New York: The Guilford Press.
- Robbins, S. P. and Judge, T. A. (2012). Emotions and moods. In Robbins, S. P. and Judge, T. A., editors, *Organizational Behavior 15th Edition*. Prentice Hall.
- Shupp, R., Sheremeta, R. M., Schmidt, D., and Walker, J. (2013). Resource allocation contests: Experimental evidence. *Journal of Economic Psychology*, 39:257–267.
- Solow, J. L. and Kirkwood, N. (2002). Group identity and gender in public goods experiments. *Journal of Economic Behavior & Organization*, 48(4):403–412.
- Sonnemans, J., Schram, A., and Offerman, T. (1998). Public good provision and public bad prevention: The effect of framing. *Journal of Economic Behavior & Organization*, 34(1):143–161.
- Stanley, T. D. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives*, 15(3):131–150.
- Stanley, T. D. and Doucouliagos, H. (2017). Neither fixed nor random: Weighted least squares meta-regression. *Research Synthesis Methods*, 8(1):19–42.
- Stanton, S. J., Reeck, C., Huettel, S. A., and LaBar, K. S. (2014). Effects of induced moods on economic choices. *Judgment and Decision Making*, 9(2):167–175.

- Sugden, R. (2000). Team preferences. *Economics & Philosophy*, 16(2):175–204.
- Sugden, R. (2003). The logic of team reasoning. *Philosophical Explorations*, 6(3):165–181.
- Tavoni, A., Dannenberg, A., Kallis, G., and Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences*, 108(29):11825–11829.
- Treffers, T., Koellinger, P. D., and Picot, A. (2016). Do affective states influence risk preferences? *Schmalenbach Business Review*, 17(3-4):309–335.
- Unger, A. and Stahlberg, D. (2011). Ego-depletion and risk behavior: Too exhausted to take a risk. *Social Psychology*, 42(1):28–38.
- Västfjäll, D. (2002). Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae*, Special issue 2001-2002:173–211.
- Webb, T. L., Sheeran, P., Totterdell, P., Miles, E., Mansell, W., and Baker, S. (2012). Using implementation intentions to overcome the effect of mood on risky behaviour. *British Journal of Social Psychology*, 51(2):330–345.
- Weber, E. U. and Milliman, R. A. (1997). Perceived risk attitudes: Relating risk perception to risky choice. *Management Science*, 43(2):123–144.
- Westermann, R., Spies, K., Stahl, G., and Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, 26(4):557–580.
- Yip, J. A. and Côté, S. (2013). The emotionally intelligent decision maker: Emotion-understanding ability reduces the effect of incidental anxiety on risk taking. *Psychological Science*, 24(1):48–55.
- Yuen, K. S. and Lee, T. M. (2003). Could mood state affect risk-taking decisions? *Journal of Affective Disorders*, 75(1):11–18.

- Zeelenberg, M., Nelissen, R. M., Breugelmans, S. M., and Pieters, R. (2008). On emotion specificity in decision making: Why feeling is for doing. *Judgment and Decision Making*, 3(1):18–27.
- Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics*, 6(3):299–310.
- Zhao, J. (2006). The effects of induced positive and negative emotions on risky decision making. In *Talk presented at the 28th Annual Psychological Society of Ireland Student Congress, Maynooth, Ireland*.