Doctoral programme in Economics and Management

The effect of Trust on Innovation: A Behavioral Investigation

a dissertation submitted in partial fulfillment of the requirements for the Doctoral degree (Ph.D.) doctoral programme in Economics and Management

Andrea
Pradelli
Academic Year 2023/2024

Supervisor– Matteo Ploner

Doctoral Committee

– Prof. Sandro Casal (Università di Trento)

– Prof. Azzurra Morreale (Lappeenranta University of Technology, Finland)

– Prof. Luca Panaccione (Università degli Studi di Roma La Sapienza)

Abstract

Trust and innovation are two multifaceted concepts, whose relevance in social sciences cannot be underestimated. Rousseau et al. (1998) define trust as "*a psychological state comprising the intention to accept vulnerability based upon the positive expectations of the intentions or behaviour of another*". Instead, innovation is defined as "*the implementation of a new or significantly improved product, good or service, or process, a new marketing method, or a new organisational method in business practices, workplace organisation or external relations*" or, in other words, "*invention plus exploitation*" (OECD, 2005). Even if, as we will show in the first chapter, trust and innovation have many different definitions, they share two pillars, that is, risk and positive expectations. In fact, trusting another person implies becoming vulnerable to the risk of deception or betrayal by the counterpart. Similarly, deciding to innovate rather than to continue with an established habit usually increases the risk of failure, as March (1999) shows in the exploration-exploitation dilemma. However, people decide to trust others and to innovate because they have positive expectations on, respectively, the behavior of counterpart or the likelihood of success of an innovation. Moreover, trust and innovation are characterized by ambiguity and uncertainty. The probability of trustworthy behavior by the counterpart or of the success of an innovation is mostly unknown. Especially in social dilemmas, trust is subject to strategic uncertainty related to the choice and type of the partner, while innovation is subject to uncertainty related to the success and quality of the innovation, the capacity to outperform the competition, whose strategic moves cannot be fully understood, and the potential ability to absorb innovation arising from the market.

The relevance of trust and innovation in the economics literature has increased remarkably in the last decades. One possible explanation is the ever-growing complexity of modern society, which relies on a dense network of transactions in everyday life and where knowledge is a primary source of competitive advantage. Trust has been identified as a driver of economic growth (Knack and Keefer, 1997; Zak and Knack, 2001), inclusive growth (Asongu and Gupta, 2015), government performance and better political outcomes (Knack, 2002; Bjørnskov, 2009; Rothstein and Uslaner, 2005), and competitive advantage for firms (Hunt, 2012). While already in 1942 Schumpeter defined "creative destruction" as "the essential fact about capitalism", the pivotal importance of innovation for long-term economic growth has been acknowledged by many authors, including Solow (1956), Romer (1986), Hasan and Tucci (2010) and Pradhan (2016). We live in the era of the "knowledge economy" (Drucker, 1969), that is, an economic system where knowledge-intensive activities are the basis for the production and development of goods and services. Therefore, understanding and mastering innovation is key to gain competitive advantage.

The study of trust and innovation requires cooperation between different disciplines besides economics, including sociology, biology, political sciences and psychology. We decided to adopt a behavioral approach for our research. Behavioral economics departs from the assumption that people always act with a fully rational and profit-seeking attitude.
On the contrary, economic agents can be subject to cognitive biases such as loss aversion, anchoring bias, confirmation bias or status-quo bias. These are the results of the brain's attempt to simplify information processing and decision making and lead to systematic errors and deviation from rational behavior. Moreover, people's actions can be driven by motives other than profit-seeking, such as inequity aversion, preferences for reciprocity, the desire to adhere to social norms or to signal themselves as moral and altruistic people (Evans and

Krueger, 2009). This could help explain why agents make choices that might appear as not fully rational, because they involve unnecessary risks or a deviation from a well-established habit. Behavioral economics can provide insights into the drivers that prompt agents to trust or innovate (for example, Neto, Felipe and Bento Caleiro, 2019; Burnham, 2000. In the second and third chapter, we decided to conduct experimental studies. Experiments offer many advantages compared to observational studies or surveys. The controlled environment facilitates the identification of causal effects and the distribution of information among participants. Experimental variables are less affected by social desirability bias, which is typical of survey measures. Moreover, experiments are easier to replicate once a solid design is developed.

In our thesis, innovation is envisioned as a coordination issue, whereby members of a team must collectively decide whether to innovate. In fact, the literature shows that most innovations happen within research centres or big companies (e.g., Powell, Koput and Smith-Doerr, 1996), which have well-structured teams dedicated to research and development. These teams have internal dynamics of conflict and power. Some members might be more expert, more charismatic, more competent or older than others. The degree of openness to new ideas might vary within the team. For example, more senior members might be wary of ideas coming from junior members, or they might be more skeptical of deviations from organisational habits. Proposing a new idea could thus be perceived as a dangerous move. Team members might fear to be ridiculed or even sidelined if they suggest a possible innovation, or they might fear that the firms will not share the potential benefits with them. Fear and mistrust could thus lead organizations away from innovation, which could result in a loss of profits and competitive advantage, especially in the long term.

In the second and third chapter, we measure willingness to innovate as a choice between a well-known task, which guarantees a safer payoff, and a new task, which is riskier, but potentially more rewarding. The decision is taken simultaneously by members of a dyad, and each participant's payoff depends on the partner's choice. This design is modelled after March (1991)'s exploration-exploitation dilemma, whereby organisations must choose whether to refine, improve and strengthen an existing process, product or technology, that is, to exploit, or to search for new solutions, that is, to explore. Exploitation guarantees safer and more proximate returns, while exploration entails more risks, but it could lead to higher payoff in case of success. March proves that most organisations tend to rely excessively on exploitation, which can be dangerous for their long-term profitability. In our thesis, we nested the choice between exploration and exploitation into a prisoner's dilemma setting. Subjects learn how to play a relatively easy task by playing it for multiple rounds, then they are matched with a partner. Members of the newly formed dyad must decide simultaneously whether to play another round of the well-known task or to attempt a new task. Accordingly with the exploration-exploitation dilemma, the new task is more difficult, and it yields higher returns in case of success, but lower returns in case of failure. Subjects play all tasks alone, but a bonus or a malus is added to their payoff based on the partner's choice. The payoff structure of the game is such that cooperation on innovation guarantees the highest joint payoff for the dyad, whereas defection, that is, choosing the well-known task, consistently results in a higher individual payoff. Defecting is thus a dominant strategy, and the only Nash equilibrium of the game occurs when both members of the dyad decide to defect. We believe that social trust and information about the partner's level of trust could help the dyad escape the suboptimal equilibrium and cooperate for innovation. In fact, high-trust participants might be less scared

of unilateral defection by the counterpart, instead they would be led to believe that they will be reciprocated if they decide to cooperate. Moreover, high-trust subjects could be more likely to believe that their ideas will be listened to, and benefits of innovation might be shared with them (Clegg et al., 2022). Trust could make social norms like inequity aversion or heuristics favoring fairness and social exchange more cogent, and lead individuals to believe that cooperating is the most "moral" choice. Knowing that the partner is a high-trust type could reinforce social norms about morality and induce higher expectations of reciprocity. In fact, in our thesis we find that knowledge on the partner's trust plays a bigger role than individual trust in facilitating cooperation on innovation.

The thesis is structured as follows. In the first chapter we review the literature on trust and innovation. First, we provide an overview of all possible definitions of the two concepts. Second, we outline the contributions of behavioral economics to the study of trust and innovation. Third, we clarify the relevance of these variables in the economics and social sciences literature. To conclude, we analyse the literature on the relationship between trust and innovation and identify four main strands. In general, the literature suggests a positive relationship between trust and innovation. At the societal level, trust favors innovation by facilitating knowledge sharing, inducing a reputational loss for egoistic behavior and reducing monitoring costs (e.g., Audretsch, 2019; Akçomak and ter Weel, 2009; Lyu, Ma and Zhang, 2022). Within teams, trust favors idea implementation and idea suggestion (Clegg et al., 2002), assists in the formation of social network ties for idea generation and realization (Shazi, Gillespie and Steen, 2014) and eases cooperation and conflict management (Lazány, 2018; Maccurtin, 2008). Trust facilitates cooperation between different firms for innovation, thus boosting open innovation (Brockman et al., 2018), the formation of innovation ecosystems (Steinbruch et al., 2021) and alliances for innovation (Gulati, 1998). Lastly, firms with a high-trust internal environment outperform rivals in innovation (e.g., Tsai and Ghoshal, 1998; Sankowska, 2013; Dindaroglu, 2023; Molina-Morales and Martínez-Fernández, 2010). However, some authors suggest that the effect of trust on innovation is not linear, that is, it may become weaker or even negative after a certain threshold is exceeded. In fact, excessive trust can lead to a dangerous reduction in monitoring (Langfred, 2004) and constructive task-based conflict (Bidault and Castello, 2010), and over-reliance on ties based on trust can lead firms to take unnecessary risks and lose market opportunities (Uzzi, 1996; Molina-Morales, 2011; Budunchi, 2013).

In the second chapter, we conduct an experimental study on the relationship between trust and willingness to innovates within dyads. The experiment is divided into two stages, which take place on two different days. In the first stage, Trust and Risk Attitude are elicited through the trust game and the Bomb Risk Elicitation Task (BRET). In the second stage, subjects play two rounds of a puzzle task, then they are assigned to either the Individual or Matching treatment. In the Matching treatment, they are randomly paired with a partner. Participants are then required to choose between playing an additional round of the puzzle task or playing a new task, called the Divergent Association Task (DAT). The DAT is radically different from the puzzle task, and it is riskier, but potentially more rewarding. In the Matching treatment, participants' payoff in the chosen task depends on the partner's choice, thus creating a prisoner's dilemma setting. Uncertainty in this setting originates from the type and the choice of the partner and the lack of knowledge about the alternative task. Participants choose under ambiguity rather than risk, because the probabilities about the aforementioned sources of uncertainty are either unknown or ill-defined. We define willingness to innovate as a dummy

equal to 1 if the subject chooses the "innovative" task, that is, the DAT. While cooperating on innovation yields the highest joint payoff for the dyad, defecting and choosing the puzzle task always yields a higher individual payoff. We find that participants assigned to the Matching treatment are less likely to innovate, as they perceive the risks of strategic interaction. Although no relationship between trust and willingness to innovate is found, we observe that subjects are more willing to innovate when they expect the partner to do the same. No evidence of a nonlinear effect of trust on willingness to innovate is found. We suggest that the different rate of innovation between the Matching and the Individual treatment could be explained by ambiguity aversion and strategic uncertainty, generated by the lack of knowledge about the probability of trustworthy behavior by the partner, the identity and type of the partner and the nature of the new task. Moreover, betrayal aversion and negative emotions related to a strategic interaction, such as fear, could have driven participants away from choosing to innovate.

In the third chapter we add information about the partner's level of trust to the experimental design. Participants' trust is elicited as the average between the amount transferred in the trust game and the answer to the World Values Survey's Rosenberg question. In the second stage, which is played immediately after the first stage, participants play five rounds of a puzzle task to create a lock-in effect (Foxon, 2014). Later, they are asked whether to play another round of the puzzle task or to play another puzzle task which has a different solution strategy. As in the second chapter, the new task is riskier, but potentially more rewarding. Three treatment conditions are created, whereby subjects are matched with a partner according to their respective levels of trust and they are informed on the partner's level of trust. High-trust, low-trust and mixed pairs are created. Again, in the treatments the choice between the two puzzles is incentivized with a payoff structure which creates a prisoner's dilemma, whereby cooperating on innovation maximizes joint payoff, but defection maximizes individual payoff. We find that being informed that the partner is a high-trust type has a positive and significant effect on willingness to innovate, that is, the likelihood of choosing the new task. The effect of Partner's trust on willingness to innovate is at least partially mediated by the expectation that the partner had innovated, thus suggesting that knowing that the partner is a high-trust type could have spurred beliefs about expected reciprocity and expected trustworthiness of the partner. However, other factors might come into play, including the desire to adhere to social norms and guilt aversion, that is, the desire not to disappoint the partner. Moreover, information about the partner's trust could have helped reduce the level strategic uncertainty related to the type of the partner, whose distribution would otherwise be unknown to the subject Instead, no direct effect of trust on willingness to innovate is found. Again, participants assigned to the Baseline, where no matching or strategic interaction occurs, are more likely to innovate.

Future research could include measures of ambiguity, ambiguity insensitivity, guilt aversion, strategic uncertainty, false consensus effect and ambiguity-neutral first-order beliefs about the partner's trustworthiness. These variables were not included in our study, but they might help explain some of our results or lack thereof and identify new channels that facilitate or hinder willingness to innovate and cooperate within dyads engaging in a strategic interaction. Variables measuring personal traits such as the Social Value Orientation (SVO) and the Big Five Personality Traits could be added, as they could possibly explain how participants form beliefs about other participants' choices. However, one should be wary not to exaggerate the cognitive load of the experiment, to avoid noise from boredom effect.

Moreover, in the second and third chapter the sample was limited to respectively British people and students from the University of Trento. Enlarging the sample and including participants who are employees of real firms could increase the power and realism of our experiment.

Acknowledgements

My PhD has been a long and exciting journey, which started on October 5th, 2020, the day when I received the news that I had been accepted as a doctoral candidate at the University of Trento. Incidentally, it was also my 26th birthday.

Over these years, I had the privilege to devote myself to my greatest passion: learning. I deepened my knowledge of econometrics and statistics, strengthened my weak foundations in coding, and built a solid background in behavioral and experimental economics. Most importantly, I had the opportunity to meet wonderful people who shared this experience with me, celebrating the good moments and supporting me through the difficult ones.

First of all, I want to wholeheartedly thank my supervisor, Matteo Ploner, who has always been supportive throughout these four- almost five- years. Without his constant help it would have been almost impossible for me to design and, above all, write the code for the two experiments which are the backbone of my thesis. Whenever I started thinking that I would never have managed to complete my dissertation, he always stepped in and convinced me that, despite all odds, I would have made it. Moreover, he was also forced to withstand my constant worries about climate change and the lack of snow in the Alps. I want to thank professors Luca Panaccione and Stefano Papa, the referees who evaluated my doctoral thesis. Their suggestions and comments were both interesting and useful and allowed me to make important adjustments to my thesis.

I would like to remember the lab manager Marco Tecilla, who sadly left us in 2024. His contribution was essential, as he wrote most of the oTree code that I used for my experiment. I am deeply grateful to Professor Luigi Mittone, who cooperated with Professor Ploner to supervise me in the PhD. I also thank all the CEEL staff, including the lab manager Paolo Leoni, professors Dominique Cappelletti, Sandro Casal, Lucia Savadori and Marco Faillo, as well as my colleagues Chiara Gioia, Austeja Kazemekaityte, Ruggiero Rippo, Mostafa Goudarzi, Klaudijo Klaser, Matteo Benuzzi, Azzurra Morreale, and all the other PhD students and researchers who "lived" the lab. Above all, I am indebted to my colleague Vittorio Guida, whose technical and psychological support has been constant from the first day of my PhD.

The administrative staff, especially Nicole Bertotti and Silvia Eccher, was always very professional and helpful. Moreover, I built a special bond with the reception staff, especially Gabriele and Manuela.

On the fourth floor of the DEM building, I became part of a family, with a constant meeting at 12:30 for lunch. Alice Bartolini, Nan Zhang, Stefan Lazic, Caterina Stocchi, Marco Tomasi, Marco Compagnoni, Davide Bernardini, Enrico Cristofoletti, Giulio Galdi, Fabio Cevenini, Alessandro Sulas, Laura Moritz, Sepideh Nikounejad, Arianna Cerutti, Silvia Delladio, Tatiana Cantillo, Tesfaye Salarin, Pierluigi Giardino, Revathi Palat Rao, Agueda Solis Alonso, Marco Lomuscio, Gianluca Grilli, Sara Lorenzini and many others became much more than simple friends.

My PhD journey was not without its challenges, and I am thankful to Ezio Elasdi for his professional guidance during the final years. I am also deeply grateful to my colleague Clementina Crocé, who became my closest confidant and was always by my side with warmth and kindness.

I cannot conclude without expressing my gratitude to many people outside academia. My family, from my parents to more distant relatives, has always supported me with patience and understanding, even when I could not stop complaining. I want to explicitly thank my uncles Marcello and Massimo, my aunts Adriana and Laura, my cousins Franco, Elena, Valeria, Roberto, Giulia and Giorgio, and my grandmothers Santina and Marisa. In Modena I could always count on my "historical" group of friends, which I call "the second family": Alessia Piccinini, Andrea Cavani, Chiara Zoccoli, Stefano Sirotti, Alice Valentini, Alessandro Martini, Andrea Bassoli, Anna Gheduzzi, Francesco Bastoni, Alberto Bavieri, Carlo Caterina, Ciro Sepe, Francesco De Sario, Davide De Silvestri, Edoardo Marchetti, Enrico Torlai, Giacomo Fabio, Matilde Fondriest, Andrea Pegoraro, Roberto Grani, Stefano Cordella, Michele Granata, Daniele Borghi, Riccardo Silva, Simone Pamini and Federico Zacchi. With them I had many joyful times, including dinners and holidays, and we always helped each other in difficult times. Outside of this group, I am extremely grateful to special friends such as Nikos Michas, Maria Vittoria Bevilacqua, Paolo Tardini and Monica Liliana Pradelli, and all my roommates in Trento, that is, Marco, Lorenzo, Piersebastiano, Damiano, Daniele, Luca, Sebastiano, Edoardo, Riccardo Lani and Riccardo Lolato.

Zakynthos, "the flower of the East" has a special place in my hearth, because I always spend my summer holidays there. I owe special thanks to my cousin's husband, Tasos, and all his family. Tasos introduced to Zakynthian culture, and I started attending the local dancing festivals, called "panigyri", even if my dancing skills are still very poor. On Zakynthos I almost became a local and I met wonderful people, including Marinetta and Katerina Paganopoulou, Fotini Theodosi, Konstantina and Dimitra Charta, Nikos Theodosis, Angeliki Giditsi, Nikolaos Stoufis, Konstantinos Fragkogiannis, Dionysis Pyromallis and his family, Anna Pappa, Eliza Tsoukala, Aleksandra Boncol-Spinou, Stathis Chasalaevris, his wife Tota and his son Dionysis, who sadly passed away last year.

In November 2024, I began volunteering for the *Fondazione Sant'Egidio* in Trento. It was a wonderful opportunity to meet new supportive and deeply altruistic friends, but I also came face to face with the harsh reality of homelessness in a city as prosperous as Trento.

The list of people I am grateful to could go on endlessly, but I want to end with one important reflection. Life can be hard, at times even overwhelming, but I have learned that nothing is more important than having a strong and generous network of people who remain close and supportive, even in the hardest moments. With such friends, everything becomes possible, even completing a PhD!

Table of contents

1 Trust and Innovation: a literature review

1.1 Introduction

Trust and innovation are two key concepts in economics and social sciences. Studying these variables requires a high level of interdisciplinarity, as useful tools can be obtained from management, psychology, sociology, behavioral and experimental economics and even biology and chemistry. The importance of trust and innovation in the literature has grown exponentially together with the increase in the level of complexity of our globalized society, even if both concepts appear in the literature even in the first half of the 20th century. Given the importance of both variables for economic development, a wide body of literature has examined the relationship between trust and innovation, providing important results and many insights for future research.

The complex socio-economic system arising from globalization and industrialisation has created a rich network of contacts, through which people with different skills and knowledge are brought together to create products and services that are part of our everyday life. For example, the production of a simple object like a pen can involve technicians, scientists, economists, managers and factory workers, but nobody would have the knowledge to handle the process of mass production and distribution of pens alone. Therefore, each of these interactions requires a certain degree of trust, that is, "*a psychological state comprising the intention to accept vulnerability based upon the positive expectations of the intentions or behavior of another*" (Rousseau et al., 1998). However, trust is not only required in impersonal transactions, where people do not know each other. Even inside firms, especially medium-sized and big firms, most projects are conducted by teams of people who do know each other, although they do not always classify as acquaintances. Even in this case trusting each other is essential for the success of the project. As we will discuss in this paper, the concept of trust is multifaceted, and there are many possible definitions. The role of trust in supporting economic growth is supported by a wide literature, such as Knack and Keefer (1997), Zak and Knack (2001). According to Kenneth Arrow "*Virtually every commercial transaction has within itself an element of trust, certainly any transaction conducted over a period of time. It can be plausibly argued that much of the economic backwardness in the world can be explained by the lack of mutual confidence*." (in Knack and Keefer, 1997, page 1252).

In the 21st century, our society is moving towards a development model called "*knowledge economy*" (Drucker, 1969), that is, an economic system where intellectual capital is the primary source of development and competitive advantage. In this system, innovation plays a pivotal role. For instance, according to the 2022 Global Innovation Index, calculated by the World Intellectual Property Organization (Dutta et al., 2022), the leading countries in the ranking are Switzerland, the United States, Sweden, the United Kingdom, the Netherlands, South Korea, Singapore, Germany, Finland, Denmark, China, France, and Japan. These nations have achieved either high standards of living or robust economic growth, or often a combination of both. Like trust, innovation is a multifaceted concept. Schumpeter (1942) was one of the first scholars to focus his research on innovation. He coined the term "creative destruction", that is, the process through which new goods and processes replace the old ones and create a new economic structure, to the point of being "the essential fact about capitalism" (Aghion and Howitt, 1992, p.324). The OECD (2005) defines innovation as "*the implementation of a new or significantly improved product, good or service, or process, a new marketing method, or a new organisational method in business practices, workplace organisation or external relations*" or,

in other words, "*invention plus exploitation*". The fundamental role of innovation for economic and social development has been stressed by authors such as Solow (1956), who defines technological progress as the residual in the equation linking the growth rate of output to the accumulation of capital and labor. In his model, technological process is the only possible driver of long-run economic growth. This result has been confirmed by Romer (1986) and Aghion and Howitt (1992).

Having asserted the importance of the concepts of trust and innovation, the goal of this dissertation is to investigate whether these two variables are interrelated and whether the former has a positive effect on the latter. The literature investigating this relationship is already quite extensive, and the most frequent finding is that trust fosters innovation. In particular, higher social trust favours innovation inside teams (e.g., Shazi, Gillespie and Steen, 2014), societies where interpersonal and institutional trust are higher show better innovation outcomes (e.g., Akçomak and ter Weel, 2009), reciprocal trust facilitates innovation processes that require collaboration between different firms, such as open innovation (e.g., Brockman et al., 2018), and firms working in an external and internal high-trust environment outperform other firms in innovation (e.g., Meng et al, 2019). However, some authors also found a "dark side of trust", that is, excessive level of trust within a team or over-reliance on relationships with highly trusted partners can lead to a lower level of innovation.

Our dissertation aims at investigating the impact of social and institutional trust in different teams and societies. More specifically, we will study the psychological and cultural motives that drive the choice to trust and innovate and lead high-trust teams to have a different innovation performance. To reach our goal, we decided to adopt tools from behavioral and experimental economics. In fact, behavioral economics allow to study the phenomena of trust and innovation in a context in which agents are not fully rational and do not only aim at maximizing individual gain. Experiments provide multiple advantages over surveys or observational studies, including an easier identification of the causal effect, a lower susceptibility to selection bias and a better replicability.

The goal of this paper is to review the literature on trust, innovation and the relationship between these two concepts. The first section will be focused on trust. First, we will present the most relevant definitions of trust and social capital, then we will analyse how trust is studied in the behavioral economics literature and lastly, we will focus on the importance of trust for economic growth and societal well-being. In the second section we will review the literature on innovation, following the same steps adopted for trust. The difference between innovation and creativity will also be stressed. In the third section we will discuss the literature on the relationship between trust and innovation. Finally, we will draw some conclusions and suggestions for further research and then locate our thesis within the literature we reviewed.

1.2 Trust

1.2.1 Definitions

Whenever we interact with another person, we consciously decide to place resources at the disposal of the other party, with no guarantee that they will be returned. As general as this assumption may be, there is no univocal definition of trust. In fact, we can define it according

to the object of one's trust, the mental process generating the decision to trust, the expectations we have on others and the context-dependent or innate nature of trust.

The OECD (2017) defines trust as "*A person's belief that another person or institution will act consistently with their expectations of positive behaviour*". This definition resembles that of Fukuyama (1995), according to whom trust is "*The expectation that arises in a community of a regular, honest and cooperative behaviour of other members of the community, based on a set of shared norms*" (1995, page 26). The norms cited by the author can refer to complex moral issues, like religion or justice, but also more practical matters such as professional associations' deontological codes. Coleman (1990) views trust actions as "*a subclass of actions involving a risk".* Any trust relationship involves at least two people: the trustor, that is, one that gives her trust to another person, and the trustee, that is, the person who is the object of trust. Risk arises because the trustor cannot completely monitor the actions of the trustee. By giving trust to the trustee, the trustor makes herself vulnerable to opportunistic behavior and betrayal by the trustee. In Coleman's model, one decides to give trust to another person if the expected benefits from this choice outweigh the expected losses.

Poppo, Zhou and Li (2015) underline the difference between *calculative* and *relational* trust. The former is based on a system of prizes and punishments, whereby one believes that the benefits of honoring a pact and the costs of not honoring it outweigh the benefits of opportunistic behavior and the costs of trustworthy behavior. Relational trust, instead, is based on past behavior of the counterpart or on a sense of common identity. Another more elaborate classification comes from Markova, Linell and Gillespie (2007). The first distinction is between *primary trust,* that is, instinctive, natural trust, like that of a child towards her mother, and *reflective trust,* that is, trust originating from a rational decision following a cost-benefit analysis. The second distinction is between *micro-social,* related to single individuals, and *macro-social trust*, directed at the whole society or some groups, for instance institutions. In the literature these two concepts are often labelled as *particularized* and *generalized trust*. If we combine these four categories of trust, we can get a more precise classification. Macro-social trust can be *a priori generalized* if it is primary or *context-dependent* if it is reflective. Micro-social trust can be *basic* if it is primary or *inner dialogicality* if reflective.

Another distinction could regard the object of trust. We could, for example, distinguish between trust in others, that is, *interpersonal trust* or *social trust,* and trust in institutions, that is, *institutional trust* or *political trust* (Murtin et al., 2018). Institutional trust can be pointed towards different institutions, for example, the local or national government, the police, international organisations etc. Other objects of trust could be the economy (i.e., *consumer confidence index* or *business confidence index*), corporations, the future etc.

For our study, we will adopt Rousseau et al.'s (1998) definition: "*Trust is a psychological state comprising the intention to accept vulnerability based upon the positive expectations of the intentions or behaviour of another*". This description includes the two fundamental elements of trust, that is, risk and a positive expectation on the counterpart. In fact, trust needs to involve the possibility for loss and regret, as the trustor cannot always monitor the trustee's actions. Trusting a partner means being subject to the risk of deception and betrayal (Delle Foglie, Papa and Spagnolo, 2025), as the partner might choose not to return one's trust, especially in strategic contexts. Often the probability of trustworthy behavior by the counterpart cannot be clearly defined, especially when little is known about the type and identity of the trustee. This

leads to strategic uncertainty, that is, "the uncertainty with respect to the purposeful behavior of other players in an interactive decision situation" (Bruttel et al., 2023) and ambiguity, that is, "uncertainty about probability, created by missing information that is relevant and could be known" (Frisch and Baron, 1988). Ambiguity and strategic uncertainty differ from risk, because under risk the probabilities of different outcomes are known. Depending on the situation, an individual deciding whether to trust a partner can find herself acting under strategic uncertainty, ambiguity or risk, and might prefer a safer option rather than placing themselves in a vulnerable position. However, positive expectations about the trustworthiness of the trustee can convince the trustor to trust, thus allowing the transaction to take place.

A concept which is close to trust is *social capital*. Like trust, social capital has many possible definitions. One of the first authors to describe this concept was Bourdieu (1986), who defines social capital as "*the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance or recognition*". This definition mainly focuses on the concepts of "resource" and "possess". More recent definitions adopt the point of view of an individual who is involved in social relations to reach her goals. The OECD defines social capital as "*networks together with shared norms, values and understandings that facilitate co-operation within or among groups*", while the World Bank defines it as "*the institutions, relationships, and norms that shape the quality and quantity of a society's social interactions*. […] *Social capital is not just the sum of the institutions which underpin a society – it is the glue that holds them together* "(Grootaert and Van Bastelar, 2002). Both definitions stress the importance of shared norms to enable social interactions. The World Bank gets to the point of characterizing social capital as the glue that holds a society's institutions together. These networks and common rules generate trust and allow members of a community to work together, generating a positive impact on well-being and economic growth. For example, Putnam (1993) and Fukuyama (1995) show that a higher level of social capital accumulated historically in Northern Italy rather than in Southern Italy led to a greater development of the former with respect to the latter, despite the two being unified under the same set of rules since 1861. Trust and social capital are two different, but related concepts. In fact, trust in others and trust in institutions is more likely to emerge in societies with a high level of social capital, as the set of norms and networks that constitute a society makes it easy to see other individuals as trustworthy. We could thus state that trust is one of the components of social capital, if not one of the most important.

1.2.2 Trust in behavioral economics

Scholars have attempted to explain why individuals decide to trust other people, even when this implies foregoing part of their personal gain. For example, Evans and Krueger (2009) provide evidence of other-regarding behavior in strategic games such as the trust game, the prisoner's dilemma, the public goods game. These games, albeit with different settings, provide players with the alternative of cooperating with partners for the common good or defecting for personal gain. The equilibria, based on the assumptions of neoclassical economics, should correspond to a set of strategies where all participants defect. However, evidence from many studies (e.g., McCabe and Smith, 2000; Dunning, Fetchenhauer and Schlösser, 2012; Zak, 2005) shows that a significant share of participants decide to cooperate.

There are many possible explanations. Social preference theories argue that individuals, while still being utility-maximisers, treat the payoff of others as positive utilities. These preferences

include inequality-aversion, whereby individuals desire their payoff to be reasonably similar to those of other participants, and altruism, whereby individuals are willing to suffer a loss to themselves to benefit another person, without expecting any reward. Inequality aversion and altruism are unconditional motives of cooperation, that is, they do not depend on beliefs about the partner and partner's actions. However, Cox (2004) and Di Bartolomeo and Papa (2016a) have proven that, besides inequality aversion and altruism, participants might be induced to cooperate due to trust and positive reciprocity, that is, participants' desire to assist people who have treated them kindly. Both studies developed a triadic design comparing participants' behavior in a trust game and two types of dictator games. In the trust game (Berg et al., 1995) players are divided into pairs and given a fixed endowment, which is usually 10 ECU (Experimental Currency Units). The first mover decides how much of her endowment she wishes to transfer to the second mover. This amount, which is a measure of trust, is tripled, and the second mover can decide how much of her new endowment to return to the first mover, thus providing a measure of trustworthiness. A more detailed description will be provided in the second and third chapter. In the dictator game (Kahneman, Knetsch and Thaler, 1986), the proposer has a fixed endowment and can decide how much to transfer to the recipient, who cannot make any choice. Rates of cooperation are higher in the trust game, where beliefs about the partner's trustworthiness and the trustee's desire to reward the trustor for being kind to her can influence participants' choices. While Cox finds evidence of both trust and positive reciprocity, Di Bartolomeo and Papa find clear evidence of the role of trust, while positive reciprocity only plays a role when initial endowment inequality between the trustor and the trustee is low.

Norm-based theories, instead, assume that choices are influenced by social norms. For example, referring to another player as "partner" rather than "opponent" enhances trust and trustworthiness, while inducing players to think about money enhances selfishness (Burnham, 2000). Heymann and Ariely (2006) introduce the difference between monetary markets, which encourage quid-pro-quo behavior, and social markets, which encourage acts of reciprocity without expecting something in return. The first type of market mainly induces calculus-based trust, motivated by the fear of punishment and the hope of a reward from reciprocity, while the latter induces identification-based trust, motivated by empathy towards the counterparty. Evidence from self-reported data suggest that people perceive games involving trust and reciprocity as social. In fact, authors like Krueger et al. (2008) and Malhotra (2004) find that players in these games interpret either their behavior or the behavior of others in moral terms. If one believes that she or their opponents decide whether to trust the counterpart based on moral standards, then cooperation becomes both rational and profitable. In both experiment and everyday social interaction, people mostly want to avoid letting the counterpart down. This mechanism, called guilt aversion, often leads to cooperative behavior. Battigalli and Dufwenberg (2007) identify two types of guilt aversion. Simple guilt consists of the belief that one's actions are leaving the partner with a lower payoff than expected, while guilt from blame implies that people interiorize the belief that the partner thinks that they are willingly acting to diminish her payoff. Both channels can induce people to cooperate in social dilemmas.

People might be subject to self-similar reasoning (e.g., Li, Turmunkh and Wakker, 2019) and form their expectations about the other's choice based on their own choice. Therefore, trustworthy people might think that other people are as trustworthy as they are and thus be

induced to trust, while the opposite would happen for selfish people. Moreover, trustworthy people might want to signal their type by trusting the partner. However, the distribution of types is mostly unknown, and the belief that others are similar to self can be false. In particular, trustworthy people tend to overestimate the prevalence of trustworthy types in the population, and this explains why trusting people often lose money. This is known as the false consensus effect, that is, a cognitive bias that leads individuals to "see their own behavioral choices and judgments as relatively common and appropriate to existing circumstances while viewing alternative responses as uncommon, deviant or inappropriate" (Ross, Greene and House, 1977, p. 280). Delle Foglie, Papa and Spagnolo (2025) find that selfish people are more likely to believe that others are similar to self. Even though both selfish and trustworthy people are moral, selfish people believe that they can maximize their economic gain while minimizing non-monetary disutility from opportunistic behavior, as they think other people are also selfish. This mechanism, called self-serving belief manipulation, leads people away from cooperative behavior.

Dunning, Fetchenhauer and Schlösser (2012) find that the decision to trust others does not only depend on economic dynamics, but also on emotional and social factors. The authors use the trust game to measure players' trust and trustworthiness. They find that participants do trust even when their expectation on the partner's trustworthiness is below their general tolerance for risk, while there is no evidence that the choice to trust is motivated by the desire to enlarge the pie to share with the partner. Instead, trust seems to be mostly influenced by the emotions that participants feel regarding the decision to trust. More specifically, immediate emotions, that is, emotions attached to the act of trust as such, are more likely to affect the choice of trusting a third party rather than anticipated emotions, that is, those attached to how participant would feel after knowing the outcome of their decision. For example, people would like to perceive themselves as following social norms or to signal to themselves and others that they are altruistic and moral. Moreover, people show a significantly higher propensity to trust when they are in a real, ongoing "relationship" with another person rather than when this relationship is only hypothetical. This can be explained through evolutionary survival strategy or cultural factors which induce subjects to trust other people when they are in any type of relationship. Delle Foglie, Papa and Spagnolo (2005) find that participants engaging in a trust game paired with a machine rather than a human being experienced more anger towards the partner related to an anticipated or real betrayal. This could help explain why participants paired with a human partner were less likely to trust in the author's experiment. Anger is measured together with other emotions in a synthetic indicator called Valence, which includes happiness, sadness, surprise, scare, disgust and contempt. Data are generated through the Noldus' FaceReader software. Di Bartolomeo and Papa (2016b) find that subjects who practice short-term meditation before playing a trust game transfer higher amounts to their partner. The authors find two possible channels to explain this result. First, meditation reduces cognitive rigidities, that is, resistance to changes in beliefs and attitudes and the perseverance in the use of mental or behavioural sets, and the diffidence and sense of competition between individuals. Second, the rituals induced by meditation create a sense of belonging to the same group, thus supporting cooperative behaviour.

Zak (2006) uses neuroeconomics, that is, a new transdisciplinary field that employs measurement techniques of neuroscience to understand how people make economic

decisions. The goal is to find whether the decision to trust is influenced by physiological, biological factors that belong to the sphere of the subconscious. Subjects are randomized to play either a one-shot trust game or a control game in which DM1 pulls a ping-pong ball numbered from 0 to 10 from an urn and this amount is subtracted from her endowment and sent to DM2, who can then decide how much to return. Moreover, participants will be taken four tubes of blood from them to search for possible spikes in a hormone called oxytocin (OT). This hormone was found to be responsible for the high level of prosociality in some species of rodents and it is released by mothers during childbirth, inducing them to attach to children and vice-versa. The authors find that OT levels in players 2 that receive intentional trust is higher than in those who receive an amount drawn randomly from the urn, even if the average amount received is the same. Moreover, players 2 who receive intentional trust and have spikes in oxytocin are more likely to return larger amounts to player 1. In fact, OT rises when player 2 receives a signal of trust from the partner, and this creates a sense of attachment and empathy to player 1 which turns into a higher return of money. Instead, there is no relationship between basal OT levels of players 1 and trust: in fact, if the basal level of hormones induced people to be more likely to give away resources, they would have been easier target for predation, and this characteristic would have been evolutionarily weak. Instead, it is the sudden spike, not longer than 3-5 minutes, induced by a social stimulus that enhances trustworthiness in player 2.

McCabe et al. (2001) employs a binary-choice version of the trust game inside a magnetic resonance imaging (MRI) scanner. They measure neural activity in the medial prefrontal cortex (BA10), which is shown to be associated with the "theory of the mind", that is, the ability of humans older than four years to anticipate what others will do by putting themselves in someone else's shoes. They randomize subjects to play the trust game either with a partner or with a computer. They find that regional neural activity in BA10 is higher in participants who play with an actual partner, and it is higher in subjects that decide to cooperate. This shows that one needs a more intense prefrontal activity in a strategic situation when she needs to anticipate what the other player will do, rather than when she plays against a computer, otherwise she would simply play the Nash equilibrium strategy and not cooperate.

Similarly, Rilling et al. (2002) find that women who play a binary-choice sequential prisoner's dilemma game with a partner have greater neural activation in dopamine-innervated midbrain regions, frontal regions associated with attention and error monitoring, as well as frontal regions that process emotions. These areas rich in dopamine receptors are the most active during rewarding behaviors.

In Kosfeld et al. (2005) subjects play a round of the trust game, then they receive either 24IU of intranasal oxytocin or a placebo. After waiting 50 minutes for the drug to take effect, subjects play four rounds of the trust game, being rematched with a different player in each round. The authors find that players 1 who receive the drug are more trusting than those who receive the placebo, as OT reduces the anxiety associate with trusting a stranger. Instead, receiving exogenous OT does not have any effect on players 2 who receive a positive amount from player 1, because being trusted itself produces an endogenous OT release, generated from the body, and additional exogenous OT does not have any physiological effect.

To conclude, there is substantial evidence that trusting behavior does happen, even in strategic settings when defecting is the dominant strategy. People are willing to forego potential gains originated from self-regarding or opportunistic behavior. This might be explained by many factors, including trust and reciprocity, other-regarding preferences, social norms, emotions or biological and chemical processes. As we will explain in the following subsection, trusting behaviors benefits society through many different channels, generating growth, facilitating investments, competitive advantage and better political outcomes.

1.2.3 Trust and economic and social development

The relationship between trust and economic and social development has been the object of a wide body of literature. Trust has been described as "*the first task of the ego*" (Erikson, 1950) because it is a fundamental tool to facilitate any type of social interaction.

Knack and Keefer (1997) find that trust and civic cooperation are positively correlated with economic growth. The trust index is obtained from answers to *World Value Survey* question: "*Generally speaking, would you think that most people can be trusted, or that you can't be too careful in dealing with people?"*, where respondents could reply with a score from 0 to 10. This is a measure of generalized trust, that is, trust in people that you do not know. Besides the relationship between trust and growth, the authors also find a positive correlation between trust and an index of "perceived government performance", "quality of bureaucracy" and "contract enforceability", and a negative correlation between trust and "government corruption", "risk of repudiation of government contracts", "risk of expropriation of assets". Moreover, they suggest that the impact of trust on growth is bigger in poor countries, where contract enforcement is not guaranteed, and the financial system is weaker. However, given that work specialization leads to more contacts between distant people, in rich countries the negative effect of transaction costs on trust is higher than in poor countries.

Zak and Knack (2001) find a positive correlation between trust and investment and trust and growth. The authors develop a general equilibrium model in which consumers interact with brokers, which are the only intermediaries to access capital markets. Consumers and brokers are identified by a type describing their characteristics and trustworthiness, but consumers only know the distribution of types of brokers and not the type of broker they are associated with. In fact, in real-life situations it is difficult to get to know even the distribution of types of partners (e.g., Li, Turmunkh and Wakker, 2019). Only the broker knows the yield of the investment, therefore a moral hazard problem arises. The consumer needs to decide how much time to dedicate to monitoring the broker, knowing that there exists an opportunity cost, as this time could be dedicated to productive activities. The amount of time the consumer dedicates to monitoring is a measure of trust. Brokers can be disincentivized from opportunistic behavior by sanctions from formal institutions, but also by the risk of sanctions from informal institutions, as ostracism from a social group. A higher level of trust is associated with higher investments and growth. Although, as expected, per-capita income is negatively correlated to growth (i.e., the advantage of backwardness), the authors find that income does not grow more for poor countries with a very low level of trust. This provides evidence of a poverty trap caused by mistrust. In fact, below a certain threshold of trust investments will not generate growth, while if that threshold is overcome the country will grow at high rates. Institutions must intervene to generate trust to break the poverty trap and allow poor countries to benefit from the advantage of backwardness.

Asongu and Gupta (2015) find a positive and significant relationship between trust and the index "quality of growth", which, besides traditional components of growth, also contains social dimensions like health and performance at school. This means that trust also favors inclusive growth. Hunt (2012) shows that firms located in an area where moral codes promoting trust are more widespread have a competitive advantage, as they have fewer monitoring costs and a reputation for trustworthiness, which facilitates strategical alliances with external firms. Chami and Fullenkamp (2002) show that promoting a trusting environment between employees of a firm is an effective tool to reduce agency problems, which arise each time someone needs to rely on the work and expertise of another person, who could use her position to behave opportunistically.

Putnam (1993) shows that Northern and Central Italy, where social capital is stronger and trust is more widespread, has reached better economic, social and political results than Southern Italy, despite the two being part of the same unified and rather centralised nation-state since 1861. Putnam identifies three ways in which social capital causes better political outcomes. In societies where trust is more widespread, monitoring of politicians and expectations about them discourages corruption and bribery. Secondly, social capital discourages political polarization, thus facilitating agreements between rival political parties and constructive criticism. Thirdly, high levels of social capital lead to higher innovation and flexibility in politics. In fact, in less trusting societies élites are often status-quo oriented and tend to focus on short-term issues, while innovative politics would be regarded by voters as not credible or aimed at favoring private interests.

Knack (2002) and Bjørnskov (2009) confirm the positive relationship between trust, social capital and government performance, while Rothstein and Uslaner (2005) and Morris and Klesner (2010) describe the positive relationship between mistrust and corruption. Rothstein and Uslaner find that generalized mistrust is the channel through which inequality causes corruption and negative political outcomes, while Morris and Klesner find a vicious cycle correlating mistrust and corruption in Mexico.

Therefore, trust is shown to have a significant role in promoting societal development or lack thereof. It is a factor that policymakers need to consider when envisioning growth-enhancing policies. It is difficult to increase trust in low-trust societies, however, if this result is achieved, a strong base for economic and social development can be created.

1.3 Innovation

1.3.1 Definitions

Like trust, innovation is also a multifaceted concept, and it has many definitions. The term "innovation" was rarely used until the twentieth century, and its etymology traces back to the late Latin verb *innovare*, meaning "altering the order of established things to make new things". One of the earliest and most famous definitions of innovation comes from Schumpeter (1934), who labels it as "*the carrying out of new combinations" that include "the introduction of new goods, ... new methods of production, ... the opening of new markets, ... the conquest of new sources of supply ... and the carrying out of a new organization of any industry*". Schumpeter refers to the concept of "creative destruction", whereby new innovations

continuously revolutionise the economic structure, destroying old jobs and wealth, but at the same time creating new opportunities for occupation and wealth accumulation, until an innovation restarts the process again. According to Schumpeter, this mechanism is the basis of capitalism: for example, railways overcame the need for horses and buggies, but later, at least in the United States, they themselves became outdated due to the competition of cars and airplanes. Nowadays, the climate crisis and the need to reduce emissions have given new rise to railways, especially high-speed lines, at the expense of airplanes, and electric cars are replacing diesel and gasoline-powered cars.

The OECD (2005) defines innovation as "invention plus exploitation", that is, "*the implementation of a new or significantly improved product, good or service, or process, a new marketing method, or a new organisational method in business practices, workplace organisation or external relations*". This definition has been widened by Mulgary and Albury (2003), who see innovation as "*the creation and implementation of new processes, products, services and methods of delivery which result in significant improvements in outcomes, efficiency, effectiveness or quality".* As we will further explain in the following paragraphs, innovation is seen as something more than a creative idea or an invention. In fact, it must result in something for which consumers can pay a price (Neto, Felipe and Bento Caleiro, 2019). Many definitions highlight the role of innovation as a process in which ideas are not only generated or adopted from other sectors or organisations, but they are also transformed into value (Confederation of British Industry (CBI)/QUINETIQ, 2008; Australian National Audit Office, 2009; Damanpour and Schneider, 2009; NESTA, 2012; etc.). Innovation is considered as a process that follows invention in time. In fact, invention is the creative act, while innovation is the first or early employment of an idea by an organization or by a set of organizations with similar goals (Becker and Whisler, 1967, p. 463).

There are many possible types of innovation. For example, Herbig and McCarthy (1993) distinguish between the following:

- Continuous innovation: The least disruptive type of innovation, which involves the introduction of a product with minor technological advantages, for example product line extensions, new sizes or new flavors.
- Dynamically continuous innovation: It is more disruptive than continuous innovation, but it does not alter behavioral patterns. One can thus create a product which is in line with the most modern technologies but still has the same basic function. An example is the electric toothbrush.
- Radical innovation: It is the riskiest type of innovation, and it requires a fundamental change in technology which can create new markets or diffuse broadly throughout the industrial base. It leads to an unprecedented change in behavioral patterns with a high societal impact. The personal computer or atomic energy can be some examples.
- Process innovation: It is an innovation that happens withing the process of production or commercialization of an existing good or service. It creates higher productivity, lower costs and higher quality. Process innovations happen constantly in every industry.

The simplest definition of innovation comes from Costello and Prohaska (2013), who define innovation as simply "*doing something different*". The focus on newness is also found in De

Vires, Bekkers and Tummers (2014), "[innovation is] *the introduction of new elements into a service-new knowledge, new organization, new management/skills*" and Evers, Ewert and Brandsen (2014), "*Innovations are in a significant way new and disruptive towards the routine and structure prevailing".* This last definition recalls March's (1991) *exploration-exploitation dilemma,* which arises whenever firms are confronted with a choice with uncertain payoff and incomplete knowledge. Organizations can either decide to refine and improve existing technologies and processes, that is, to exploit, or to search for new solutions, that is, to explore. Exploitation guarantees certain and proximate, but lower returns, while exploration can lead to more uncertain and delayed, but possibly higher returns. Organisations usually prefer exploitation, but this could lead to lower profits in the long term. Positive expectations on the success of the new process or product can facilitate the choice of exploration. Although the definitions of exploration and innovation do not coincide, a certain degree of exploration is needed to pursue innovation. Moreover, it should be noted that innovation shares the two building blocks of trust, that is, risk and a positive expectation. In fact, most innovations are costly and time-consuming, thus they imply a greater risk of failure with respect to well-established habits (Costello & Prohaska, 2013). Innovation is also subject to uncertainty. In fact, organizations do not know in advance whether they will be able to innovate, the quality of their innovation and their capacity to absorb innovation from the market. Moreover, it is difficult to understand competitors' strategic moves and to determine the probability to outperform them in innovation. However, individuals and organisations decide to defy the risk and uncertainty because they have positive expectations about the probability of gaining benefits from innovation.

Some authors underline that innovation can also have consequences beyond firms. In fact, it can also bring changes in behavior, generating a new mindset and way of thinking focused beyond the present and into the future (Kuczmarksi, 2003). To conclude, Baregheh, Rowley and Sambrook (2009) envision a new, multi-layer definition of innovation as: "[..]*the multi-stage process whereby organizations transform ideas into new/improved products, service or processes, in order to advance, compete and differentiate themselves successfully in their marketplace*". First, the authors underline that innovation is not a discrete act, instead it is a process comprising multiple stages. Second, ideas need to be used and transformed into new, improved and changed products or services, for example, for new markets. Thirdly, the aim of innovation is "successfully advancing" and "competing and differentiating" in the marketplace. This reflects both the strategic aim of innovation and the potentially diverse contexts in which it takes place.

1.3.2 Creativity and innovation

Creativity is a concept that is often associated with innovation, and in fact a wide body of literature suggests that it is one of the main determinants of innovation. It is defined as the ability or process to create something *novel* (original, unexpected) and *appropriate* (useful, adaptive concerning task constraints) in a social context (Sternberg and Lubart, 1999). From this definition one can understand that there are common traits with innovation, but the two concepts do not coincide. It is essential to clarify the difference between creativity and innovation to avoid confusion between the two concepts in the second and third chapter of the dissertation, where the dependent variable will be willingness to innovate, that is, participants' eagerness to attempt something new and begin the innovation process. At the same time, especially in the second chapter, the new and innovative task requires some degree of creativity.

Sarooghi, Libaers and Burkemper (2015), Yusuf (2009), and Neto, Felipe and Bento Caleiro (2019) propose that creativity is among the raw materials of innovation, establishing therefore a positive correlation between the two variables. However, to turn creativity into innovation social, cultural, organizational, institutional and psychological factors must come into play. Sarooghi, Libaers and Burkemper (2015) find a positive relationship between creativity and innovation. The authors maintain that the conversion from creative ideas to innovation implies two different and conflicting processes. The first is idea generation, which requires explorative activities, that is, experimentation, disrupting established routines and challenging common assumptions. The second is idea implementation, which requires exploitative activities, such as routine execution, developing a process which is efficient and goal oriented. These two activities compete for scarce resources and may even inhibit each other. Ambidexterity is defined as the ability to manage the conflict between idea generation and idea implementation and find a balance between the two activities. The authors find that the correlation between creativity and innovation is stronger for individuals than for teams or organizations, because the former, despite having fewer cognitive, intellectual and skill-based resources than firms, can more easily achieve ambidexterity. Therefore, if firms want to improve their ability to turn creative ideas into innovation, they need to identify and properly incentivize ambidextrous individuals. The effect of creativity on innovation is stronger for large firms than for SMEs, because the former have more resources to handle ambidexterity, and it is stronger for process innovation than for product innovation, as the integration of knowledge from several departments appears to be easier in the former case. Surprisingly, the effect is bigger for low-tech industries, probably because most innovations in such firms are incremental, thus easier to convert from creative ideas into marketable innovations. Collectivist cultures seem to have a greater success rate in converting creative ideas into innovation, while moderate levels of risk-taking and an ability to overcome resistance to change are critical to generate creative ideas and turn them into innovations.

Yusuf et al. (2009) maintain that innovation springs from the creative application of knowledge. Although knowledge is the raw material of innovation, the creative act is the basis of innovation. A society can enhance creativity by attaching high value to learning, physical and mental health and creating *wikicapital*, that is, the capital arising from networks of people who share knowledge and bring together diverse talents. Creative societies can become innovative if they have a high tolerance of risk-taking, rewards for commercially viable innovations (Intellectual Property Rights), they accept that so-called "good inequality", can accommodate wide income differences by reinforcing incentives that attach value to singular achievements, they are open to domestic and foreign competition and have big cities with many connections and actors working together. Public and private spending in R&D are also essential. While large firms have more resources to make the transition from creative ideas to marketable products, SMEs can also reach the same goal by cooperating with large firms.

As we will see in the next paragraph, Neto, Felipe and Bento Caleiro (2019) also maintain that creativity is the "raw material" of innovation and it is not exclusively an innate ability, but it can be developed by repeated exposure to an environment that provides stimuli to "see things differently". Creativity can be turned into innovation through a neurological process whereby people transform unconscious stimuli into marketable products and services through a

conscious process that involves the control of fear and the ability to communicate the idea to colleagues and potential customers.

To conclude, the literature that we presented shows that creativity and innovation do not coincide, as innovation requires the existence of something for which consumers can pay a price. However, creativity seems to be strongly and positively correlated with innovation, to the point of being one of its "raw materials".

1.3.3 Innovation in behavioral economics

Innovation has been studied in behavioral economics, as it is the result of a choice made by individuals or teams who are likely to deviate from the *homo oeconomicus* assumptions and are motivated by factors other than rational profit-seeking.

Morrison and Pots (2008) introduce a new framework beyond choice under perfect information and choice under uncertainty, that is, choice under novelty. By novelty the authors identify a new idea, either coming from within or outside the firm, which changes the connectivity or dimensionality of the world, making the latter more open, conceptually larger and immediately more complex. Under perfect information, all information is available and costless to process, therefore the *homo oeconomicus* will make no mistakes nor choices whose results can be improved upon. Innovation will thus occur instantaneously as someone stumbles upon an idea, all agents perfectly know the consequences and value of the idea, firms and consumers will immediately adapt and reconfigure their choices as the idea is "picked up and taken to the market". Under uncertainty, instead, there will be a deficit in information regarding possible outcomes, actions of others or payoffs. Applying the *homo oeconomicus* paradigm, agents will use a correct model of the world and all currently available information to form a "rational expectation" about the future state of a variable. Although they do not have access to all information, they do know how to correctly apply the available information. Therefore, errors will be made, but they will be stochastic (i.e., stochastically distributed), not systematic. If we relax the *homo oeconomicus* assumptions and allow the choice to be made by humans who are not fully rational, systematic errors will be made as agents apply wrong models and select inappropriate choice heuristics, being influenced by cognitive biases. Instead, in the case of choice under novelty, there is no information in the initial instance, therefore it is impossible to be rational. Agents will thus need to embark in a process of discovery and learning to simply acquire information on whether and how much a given novelty is worthy of attention, what its value is and what to do with it. This process can involve different decision heuristics originating from behavioral considerations. Therefore, *behavioral innovation economics* is centred around choice under novelty. It begins with the first apprehension of novelty and ends, potentially, with the adoption and retention of novelty for ongoing use. Its goal is to identify possible heuristics and biases that lead agents to depart from "perfect innovation" and fall into *innovation failures.* Behavioral biases that can hinder innovation are, for example, risk aversion, loss aversion, myopia, status-quo bias and confirmation bias (Potts, 2010). These can cause innovation failures in many ways. It is difficult to be aware of novelty, as the adult brain tends to filter out concepts that are cognitively distant from something familiar. The brain also finds it hard to understand the consequences of novelty on oneself, thus causing poor strategic responses. Selecting between new ideas requires a partially inductive procedure, which can pave the way for behavioral biases such as loss aversion, sunk cost effects and availability or representativeness bias. Moreover, the selection of ideas depends on who promotes them, so that a selection between

ideas becomes a selection between people. Individuals and organizations tend to be skeptical about ideas coming from outsiders, thus making open innovation more difficult. It is not easy to be rational about innovation, because attitude towards new ideas may be biased by the desire to display a social identity, for example dominance, leadership, submission or cooperation. Organizations do not tend to encourage the creation and development of new ideas, as this means departing from well-established habits and routines. Those who think of proposing new ideas can suffer from anxiety and the risk of social rejection. Moreover, it would be more convenient to adopt a *portfolio* of different new ideas to reduce the risk, but thinking about portfolios is hard for the human mind. Organizations also face difficulties in investing in innovation, because they can overlook difficulties due to *optimism bias* or exaggerate the risks due to loss aversion. Innovation requires a mental space for experimentation. However, this may disconnect from past knowledge and facilitate endowment effects and sunk cost bias. To conclude, the prospect of failure is hard to cope with. Sometimes organisations do not manage to recognize failures when they occur or fail to learn from them. This leads to avoidance of experimentation.

A new branch of neuroeconomics, that is, *neuroinnovation*, has arisen to investigate how innovation is generated inside the human brain. The starting point is the distinction introduced by Kahneman (2011) between two cognitive systems, *system 1* and *system 2*. System 1 is based on intuition, and it is fast, automatic, unconscious, effortless, associative, emotionally and culturally driven, difficult to control, while system 2 is based on reasoning and it is slow, serial, effortful, and deliberative. Innovation is thus seen as a transformation process, which originates from the understanding and generation of unconscious stimuli in system 1 and turns them into new products and services through a conscious process involving system 2.

Neto, Felipe and Bento Caleiro (2019) develop a model whereby the innovation process is comprised of three steps, that is, perception of market opportunities, control of fear and social intelligence. The first step for innovation is to perceive the world differently from the common sense. An image is captured by the eye and then interpreted by the brain using the most plausible explanation, which usually comes from past experiences. However, when the brain receives new information, it does not manage to interpret it correctly. This change of environment can create trouble for our brain, but the more the brain is exposed to the same information, the faster and less effortful the association of the stimulus to an explanation will be. This process is called *repetition suppression*, and it explains why perceiving things differently from others, which is a prerequisite for creativity, is not only a natural gift, but more often it is the result of frequent exposure to environmental change and information shift. However, a second obstacle arises, that is, fear. In fact, our brain is biologically wired to reject new ideas, as fear is an emotion that lives in system 1 and prepares us to hide or escape in front of a danger. In a work environment, employees will face fear of ridicule, of failure and of the unknown whenever they wish to defy a routine and communicate a new idea to their superiors. Thirdly, having overcome fear, employees will need to communicate their idea to their colleagues and potential customers in the market. This requires social intelligence, that is, *the capacity to generate power to access and have influence over people*. (Neto, Felipe and Bento Caleiro et al., 2019, p.19). Therefore, innovation is a dialectic process involving both system 1 and system 2. It can be portrayed through a model of brain learning. In such a model, one starts with *unconscious ignorance*, that is, not knowing that innovation exists, then

progresses to *conscious ignorance,* where she knows of the existence of innovation, but does not know how to innovate. The third step is *conscious knowledge,* where the subject knows what innovation is and innovates through a conscious elaboration involving system 2. The final step is *unconscious knowledge,* where innovation becomes an automatic process grounded in system 1. In fact, the process of *repetition suppression* allows some neurons to become experts in the development and implementation of innovation, turning it into an unconscious process.

1.3.4 Innovation and economic and social development

The pivotal importance of innovation for economic and social development has been widely stressed in the literature. Among the first scholars to acknowledge the role of innovation in the modern economy was the Austrian economist Alois Schumpeter, who defined the concept of *creative destruction* as the process that *"incessantly revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one*" and "*is the essential fact about capitalism*" (1942). Through creative destruction new goods, services and markets gain relevance in the economy until they replace the old ones, only to be replaced themselves by new innovations. This keeps the engine of capitalism in motion.

The Solow model (1956) is one of the most well-known models explaining economic growth. In the simplest version of the model, output is simply a function of capital and labor. Although both factors have constant returns to scale, they also have decreasing marginal returns. This means that, for example, when the stock of capital in the economy is very low, even small increases in capital will have a remarkable positive effect on output. However, when the stock of capital is already high, the same increase in capital will lead to much smaller increases in output. In the long run, the economy will converge to a steady state in which both output per capita and capital per capita are constant and the level of investment and savings per capita just equals the level of depreciation of the capital stock per capita. There will be no long-run growth of output per capita, as it would require capital per capita to increase faster than output per capita. If we introduce technological progress, which is defined as the "residual" part of output growth not attributable to capital nor labour, long-run growth is possible. Technological progress allows for a higher production with the same amount of labour and capital, thus increasing the productivity of these two factors. Therefore, in the steady state output per capita grows at a rate equal to technological progress and the economy will be in a state of balanced growth, with technological progress and population growth being the sole factors influencing the growth rate of production. This means that in the long run an economy with a high rate of technological progress will grow faster and eventually outperform other economies.

Romer (1986) departs from traditional aggregate long-run growth models, where in the long run wage rates and capital-labor rations across countries are expected to converge, reaching a steady-state with no growth. He adds knowledge, which is a product of research undertook by forward-looking and profit-maximizing agents. Although research has diminishing returns in the production of new knowledge, knowledge has increasing returns in the production of output. This means that, as knowledge increases, output can grow unbounded, and countries that are more capable of producing knowledge can exhibit persistently higher growth rates of output. Moreover, as knowledge cannot be perfectly patented or kept secret, investment in knowledge by one firm produces positive externalities for other firms. Therefore, the

equilibrium will be characterised by a positive rate of growth, as it is not optimal to stop at a steady state where knowledge is constant, and no research is undertaken.

Aghion and Howitt (1992) focus on industrial innovation to create a model of growth through creative destruction. Similarly to Romer's model, the expected growth rate depends on the amount of research in the economy. Firms are driven to invest in research because of the monopoly rents they would get if they patented an innovation. However, these monopoly rents are only enjoyed for a limited period, that is, until an innovation is patented and destroys existing rents. Moreover, the expectation of more research will increase the expected demand for skilled labor, thus increasing skilled workers' real wage. These two effects will erode monopoly rents and create a situation where research in period t is negatively correlated with expected research in period $t + 1$. In the model, society must allocate the flow of skilled labor between manufacturing and research. There is a unique stationary equilibrium with positive growth. However, other equilibria may also exist, including the possibility of the economy oscillating between two equilibria, one of which is characterised by the absence of research and growth.

Hasan and Tucci (2010) provide additional evidence that innovation leads to economic growth. In fact, they find that the ratio of R&D expenditures to GDP and the ratio of total patents granted to total R&D expenditures are positively correlated with per capita GDP. They also find that quality innovation, measured by the proportion of patents granted in the US on total patents granted by a country each year, leads to even higher growth than average innovation output. Moreover, they provide evidence that countries that increase the level of patenting activity, especially if granted in the US, show increases in the growth rate. Pradhan et al. (2016) demonstrate that the long-run economic growth of Eurozone countries is positively affected by both innovation and financial development. A more blurred view comes from Maradana et al. (2017). Studying European countries, the authors find some instances of a supply-leading effect, where innovation fosters economic growth, but also some instances of demand-following effect, whereby economic growth foster innovation. Moreover, in some cases the relationship between the two variables is either bidirectional or absent. Therefore, policymakers who want to foster innovation in their country should not only focus on innovation-enhancing policies, but they should also focus on fostering economic growth.

To conclude, evidence suggests a strong link between innovation and economic growth, and this relationship is expected to become even more important as far as society progresses towards the "*knowledge economy*" (Drucker, 1969), that is, an economic system based on intellectual capital. As shown in the literature, economies that invest more in innovative sectors will have a competitive advantage over economies specialising in low value-added sectors. The role of education and research and development will be pivotal, as well as that of a positive attitude towards risk and openness to novelties.

1.4 Trust and innovation

As we have seen, both trust and innovation have a central role in fostering economic development and have been the object of many studies. Both variables are also particularly relevant for behavioral economics. The aim of this section is to review existing studies on the relationship between trust and innovation. In the next section, we will suggest new paths for future research, especially in the sector of behavioral economics.

Studies on trust and innovation deal with the topic from different perspectives. Therefore, the literature can be divided into four strands of research.

A first strand of literature analyses the relationship between trust, social capital and innovation at the societal level. Audretsch et al. (2017) find that social trust strengthens the positive effect of tolerance on innovation. In fact, more trusting people are more likely to engage in knowledge spillovers and innovation activity, and the more they feel safe to trust others, the more they will be willing to tolerate diversity and individuality. The nature of this effect, however, is influenced by the level of economic development and the strength of formal institutions. Thompson (2018) finds that social capital stimulates innovation activities and consequently economic growth. The author assumes complementarity between intermediate goods and services; thus, monopolistically competitive and profit-seeking firms benefit from collaboration and sharing of information. As seen in Aghion and Howitt (1992), firms decide to enter a market if they expect to gain monopoly power over an innovative good or service after they manage to patent it. The existence of complementarities between intermediate goods makes cooperation and the endogenous accumulation of collective social capital possible. Therefore, not only does social capital increase innovative activities, but innovative activity itself, requiring exchanges and cooperation between monopoly-seeking firms, increases the level of social capital in the economy. Similarly, Akçomak and ter Weel (2009) demonstrate that social capital increases innovation outcome (i.e., the relative number of patent application), which in turn leads to income growth. They develop a model of production in which knowledge accumulation increases income, and the stock of social capital has a positive effect on the accumulation of knowledge. Social capital can help investors cope with the element of risk which is always present in innovation. Firstly, social capital can prevent egoistic behavior by the counterparty by inducing the fear of reputational loss. Secondly, it can allow researchers to signal the quality of their ideas by displaying an honest character. Thirdly, if the relationship between the investor and the researcher is based on trust, monitoring costs are lower, and it is easier for venture capitalists to gather information about the quality of the firm and its projects. Similar results are found by Lyu, Ma and Zhang (2022), who find that higher social trust in each region favors corporate innovation. The authors find three possible channels. First, social trust helps mitigate the agency problem in the relationship between the principal, that is, the shareholder, and the agent, that is, the manager, as it disincentivizes managerial opportunism by establishing strong social norms that cause reputational loss if they are violated. Secondly, social trust can help reduce the financial constraints on innovation. In fact, innovation is a cash-consuming activity, and it needs a high degree of financing from external investors. However, there is a remarkable information asymmetry between the investor and the firm, which the firm may exploit to divert the funding to activities unrelated to innovation. As in the previous case, social trust creates reputational losses for firms behaving opportunistically, so investors are more likely to finance innovation-oriented firms, thus reducing the financial constraint. Thirdly, social trust favors cooperation. This in turn eases the exchange of knowledge between different firms and facilitates the development of innovation. Dakhli and De Clercq (2004) find that human capital is positively correlated with innovation, while the same holds only partially for social capital. Innovation is measured as a composite variable including the number of patents in a year, the percentage of expenditure in R&D and the volume of high-tech exports relative to total exports. Generalized trust, that is, trust in others, positively affects the number of patents and the level of R&D expenditure, by facilitating extensive cooperation and a freer

exchange of information through the reduction of costly and time-consuming monitoring. Institutional trust fosters high-tech export, as it facilitates the interaction with government agencies and chambers of commerce, which promotes and stimulates the export of products with high value-added products. To conclude, a game-theoretic approach to the diffusion of innovation in a society comes from Leon, Etesami and Nagi (2022). The authors develop a coordination game played by individuals belonging to a social group, who must choose whether to adopt a product A or a new product B. They describe a limited-trust equilibrium, in which, unlike the Nash equilibrium, players not only care about their own well-being, but they are willing to sacrifice part of their short-term utility by behaving trustworthily, to benefit their neighbors and obtain returns in the long-run. As a result, when the level of trust in the social group is higher, it is more likely that the innovation prevails over the traditional product, and it does so at a faster rate. Moreover, individuals who behave trustworthily gain significantly higher payoff than those who are solely self-interested, thus making trustworthiness a rational choice.

A second strand of literature examines the effect of trust between members of a team on the latter's ability to innovate. Clegg et al. (2002) define innovation differentiating between *idea suggestion*, that is, to what extent one has suggested new work targets or objectives, and *idea implementation*, that is, to what extent have one's suggestions been implemented. They find that "*trust that heard*", that is, the feeling that someone will listen to your ideas, is positively associated with idea implementation, while *"trust that benefits",* that is, the belief that the organization will share the benefits of an innovation with the person who proposed it, only predicts idea suggestion. Similarly, Shazi, Gillespie and Steen (2014) investigate the role of trust in the formation of social network ties for idea generation and idea realisation, which are two stages of innovation. They define three dimensions of trust, that is, characteristics of the trustee that make her trustworthy: ability, benevolence, and integrity. They find that benevolence and ability are significantly and positively correlated with the formation of ties for both idea generation and idea realization, while integrity is only positively correlated with ties aiming at idea generation, and only in large research-intensive firms. Moreover, benevolence moderates the relationship between ability and tie formation for both idea generation and idea realisation, while integrity moderates the relationship between ability and tie formation for idea realisation. This means that if a partner is not considered benevolent, she will not be accepted as a partner for idea generation or realisation even if she is considered competent. The same holds with integrity for idea realisation. Lazány (2017) comes to the general conclusion that people in trusting relations are more prone to create and cooperate and are better at communication and conflict management. However, developing trust requires time and constant positive feedback from the person trusted, therefore firms must invest in creating a trusting environment to reap the positive effect of trust on innovation. This requires investing not only in trust among employees, but also between subordinates and supervisors, so that innovative ideas from new entrants can be put to work. Trust between upper echelons is also important. In fact, Maccurtain et al. (2008) show that trust between members of the Top Management Team (TMT) influences innovation indirectly. In fact, trust within the TMT is positively correlated with task reflexivity, that is, the extent to which team members collectively reflect on issues regarding the team and react accordingly, and the ability to share knowledge. However, some authors have found contradictory results on the effect of trust in a team's performance. Langfred (2004) conducts a study on self-managing teams, that is, "groups of interdependent individuals that can self-regulate on

relatively whole tasks" (Cohen, Ledford and Spreitzer, 1996; in Langfred, 2004, p. 385). He finds that groups who show higher interpersonal trust tend to reduce monitoring, as proposing to monitor other team members may be seen as a betrayal of the trusting relationship inside the team. Introducing in the study the level of individual autonomy, the author finds that teams with high individual autonomy and low monitoring have a lower performance than highly autonomous groups more dedicated to monitoring. In fact, when teams with a high level of individual autonomy also have high trust among members, they will dedicate less effort to monitoring and this will have a negative impact on performance. Another factor affecting a team's innovation performance is the presence of conflict. Jehn and Mannix (2001) identify three types of conflict. Relationship conflict regards personality clashes between members of a team, and it involves negative feelings such as tension, annoyance, frustration and irritation. Process conflict implies controversies about how the accomplishment of the task should proceed, including issues of duty and resource delegation. Task conflict involves an awareness of different opinions regarding the task itself. Unlike the first two types of conflict, a moderate amount of task conflict can be beneficial, as the discussion between different ideas can generate a superior result than a decision taken by one person alone. Other authors have tried to investigate how and whether trust influences the relationship between conflict and innovation. De Clercq et al. (2009) study the moderating effect of social capital on the conflict-innovation relationship. He investigates the role of two components of two components of social capital: social interaction, that is, the strength of the social relationships between exchange parts, and trust. While social interaction amplifies the positive effect of task conflict on innovation and exacerbates the negative effect of relationship conflict, trust is shown to have a suppressing effect on the task conflict-innovation relationship. In fact, trust hampers the ability to grasp the benefits of a constructive exchange of different opinions among team members or different units of a firm. First, as highlighted by Langfred (2004), trust reduces monitoring. Second, trust discourages the inclusion of conflicting points of view in the decision-making process, because that might be seen as a betrayal of the trusting relationship among the members of the team. Instead, trust induces conformity. Similarly, Bidault and Castello (2010) conduct a laboratory experiment where subjects are grouped in pairs and are asked to design and build an object in the most creative way possible. They find a nonlinear relationship between trust and creativity, that is, while on average more trusting pairs produced more creative objects, the effect declines or even turns negative after a threshold of trust was reached. They explain this finding by highlighting that trust reduces relationship conflict, but an excessive amount of trust hinders task-based conflicts that induce critical thinking.

A third strand of literature investigates the role of reciprocal trust in fostering innovation in situations where different firms cooperate to reach a given goal. Brockman et al. (2018) find a positive correlation between societal trust and open innovation, that is, a collaborative approach where firms integrate external knowledge and expertise into their innovation processes. In fact, societal trust acts as a substitute for formal contracts in reducing transaction costs and the risk of unintended knowledge spillovers. The effect is stronger in high-tech-intensive sectors, where it is more difficult to enforce property rights, and in countries with a weak legal system and low transparency. Moreover, the authors also find that trust increases innovative efficiency, that is, the ability of a firm to generate patents from a given level of R&D investment. Similarly, Steinbruch et al. (2021) find that trust among actors boosts the development of innovation ecosystems. Innovation ecosystems are composed of

interrelated organisations (i.e., firms, government, intermediary agents, research institutes and universities), that are connected through a main organization or technology platform with the goal of generating innovation. They consider the previously listed three dimensions of trustworthiness, that is, benevolence, ability, and integrity. The authors find that all dimensions are complementary to foster the development of innovation ecosystems. Moreover, trust is reciprocal, that is, a virtuous cycle can be created, whereby trust eases interactions between actors, generating commitment and satisfaction, and the more interactions occur, the more actors will trust each other, thus facilitating innovation. Teramoto and Jurčys (2017) maintain that facilitating the dissemination and exchange of ideas is the most important policy goal in today's networked societies, as individuals who trust each other are more likely to generate and implement creative ideas. Companies should create a trust-based ecosystem, in which different intermediaries, competing, can provide them with precious information. This will favor the dissemination of ideas. Multiple exchanges between parties will enhance the accumulation of trust, and this in turn will favor the creation of an innovation ecosystem. In fact, trusting one's neighbor (i.e., the intermediary) minimizes risk and facilitates communication through intermediaries. An essential element for the emergence of innovation ecosystems is the ability of firms to create ties with each other. Gulati (1998) finds that social capital provides firms with new opportunities and tools to utilize the financial and human capital with which they are endowed. For example, social capital originated from being part of a social network provides firms with an informational advantage that can advantage them in the formation of new alliances. Moreover, firms with more social capital will be able to attract better partners and obtain superior terms of trade. This informational advantage will thus favor the development and ultimate success of the alliance. Uzzi (1996) studied embeddedness, that is, a logic of exchange in which social ties shape actors' expectations and opportunities. In contrast with arm's-length ties, firms operating in embedded ties do not exclusively focus on immediate selfish gains, but they concentrate on cultivating a long-term cooperation which can lead to collective benefits. Embedded ties could not exist without a high degree of trust between involved firms. However, the relationship between embedded ties and a firm's performance is positive only up to a threshold point. In fact, being part of embedded ties increases a firm's chance of survival, but if it only relies on embedded ties, it risks being constrained by the network's set of rules and be blinded to opportunities arising from cooperation with firms outside the network. An optimal network should therefore comprise both embedded and arm's-length ties. Molina-Morales et al. (2011) also find a "dark side of trust". Their studies on manufacturing firms from different industrial districts in Valencia finds that, although some level of inter-organizational trust is essential for innovation in industrial districts, as it favors information exchange, overinvesting in trust can have a negative impact on innovation rates. In fact, overinvesting in trust, or investing in trusting relationships of little value for the firm, can cause the firm to misallocate precious resources or take unnecessary risks. Beyond a certain threshold, additional increases of trust bring diminishing benefits and might even decrease innovation returns. To conclude, Budunchi (2013) finds that trust plays a fundamental role in the selection of suppliers in new product development (NPD), as it reduces the costs, risks and uncertainty of collaboration. However, over-reliance on trust in selecting NPD suppliers stifles radical innovation, as it encourages firms to only search for information and skills within their existing supply base. Moreover, it favours inward looking, hampering the ability to adopt new information outside the knowledge domain.

A fourth strand of literature investigates whether firms that act in a high-trust environment have a better innovation performance. Tsai and Ghoshal (1998) find that trustworthiness, one of the components of social capital, facilitates resource combination and exchange among different business units of a company and thus favors the development of innovative new products. Watson Wyatt (2002) finds that high-trust organisations outperform low-trust organisations, because high trust improves engagement, loyalty, communication, collaborations, relationships with stakeholders and, eventually, innovation. Similar findings and explanations come from Covey (2006), who demonstrates that low trust organisations pay various sorts of "taxes", such as unnecessary redundancy and bureaucracy to ensure control and hinder deceptions, a waste of time and energy to be dedicated to internal fights instead of innovation and creativity, and fraudulent activities of employees disrupting the functioning of the organisation. Similarly, Meng et al. (2019) find that firms operating in countries with higher social trust have both higher innovation input (i.e., R&D expenditure) and innovation output (i.e., number of patents). The effect is stronger for firms with stricter financial constraints and countries with weak protection of intellectual rights. El Ghoul et al. (2023) extend on the results provided by Meng et al. and confirm that firms acting in environments with high social trust have both higher innovation input and output. In detail, they find that social trust is positively correlated with the number of patents. This holds especially for the number of patents with forward citations, that is, the effect of trust on patents is driven by high-quality innovations that are acknowledged by citations in other patents. Firms operating in high-trust countries spend more in R&D and to do so with less volatility, and to be less dependent on internally generated cash flow to finance R&D, thus more likely to receive external financing. In fact, the association between innovation investment and higher market evaluation is stronger in high-trust countries, and the same holds true for the association between innovation investment and future sales growth. Sankowska (2013) further investigates the relationship between trust and innovation in firms, providing a more detailed picture of what happens. In fact, she finds that trust fosters organisational innovativeness not only directly, but also indirectly. She provides evidence that trust facilitates the transfer and creation of knowledge within the company. In fact, in firms with high organisational trust it is easier to transfer knowledge, thus it is easier to create knowledge and finally to innovate. Dindaroglu (2023) uses three measures of innovation, that is, patent applications, trademark applications and the share of innovative sales in turnover by SMEs. While the first two indicators refer to the inventive phase of innovation, when existing technological capabilities are improved, the latter also includes the phase of successful commercialization and marketing. The author finds that the level of interpersonal trust in a region is positively correlated with the first two indicators, while it shows an inverted U-shape relationship with the share of innovative sales. The authors suggest that trust has a strong positive effect for the creation of knowledge commons, that is, the governance of shared knowledge in a community. Instead, its effect on the successful commercialization of a product is only positive until a threshold is reached. In fact, an excessive amount of trust might hinder commercialization. To conclude, Molina-Morales and Martínez-Fernández (2010) undertake an analysis of manufacturing firms in the Valencia Region, some of which belong to industrial districts. The authors find that social capital factors as social interactions, trust, shared vision, and involvement of local institutions are higher for firms belonging to districts. Even within the districts, firms that score high in these factors have better innovation outcomes. As regards trust, it favors innovation because, by acting as a mechanism of informal control and reducing the need for monitoring and the fear

of being exploited by the counterpart, it facilitates social exchange, even when the latter involves unpriceable resources that are nevertheless important for competitiveness. Being perceived as trustworthy, consequently, makes firms more likely to be chosen as exchange partners. Therefore, evidence seems to suggest that firms operating in a high-trust internal and external environment tend to have a better innovation performance.

To conclude, the literature mostly points to the existence of at least a correlation between trust and innovation. There are many channels through which this correlation can manifest itself. Trust can facilitate interactions among members of a team or an organisation, as it helps reduce transaction and monitoring costs and fear of being ridiculed. This allows firms characterised by an internal and external trusting environment to have a competitive advantage in both innovation and profitability, and it facilitates cooperation with other firms in the innovation process. Even at the societal level, communities with a higher level of institutional and interpersonal trust, as well as higher social capital, are shown to achieve a better innovation performance. However, trust also seems to have a "dark side". In fact, interpersonal and interorganizational relationships characterised by high trust can lead to a lower level of monitoring and constructive task conflict. Moreover, organisations which rely solely on relationships with highly trusted partners risk losing possible market opportunities. Therefore, it seems that trust has a positive effect on innovation, however, this effect might be nonlinear, that is, after a certain threshold of trust the impact on innovation becomes less significant or even negative. In the following section we will suggest possible new research paths and highlight research gaps. Moreover, we will locate our thesis within the existing literature on trust and innovation.

1.5 Discussion and conclusion

This chapter described the state of the art of research about the concepts of trust and innovation and the relationship between these two variables. We first went through the definitions of trust and innovation, which are both multifaceted concepts which do not have a univocal definition. We underlined that trust and innovation share the same building blocks, that is, risk and a positive expectation for the future. Later, we stressed the relevance of trust and innovation in behavioral economics. In fact, both concepts are well suited to be analysed in a context where agents are not only profit-maximisers and psychological, biological and social motives are in play. As regards trust, evidence shows that, although in many social dilemmas defecting is the dominant strategy, people do display other-regarding behavior. Different theories were advanced to explain this phenomenon, including social preference, norm-based and dispositional theories, or even theories involving emotions or chemical and biological factors. Innovation is also studied in behavioral economics. Examples include new frameworks such as choice under novelty and neuroinnovation, which both explain the psychological and biological motives that can lead to successful innovation or innovation failures. It is then of paramount importance to distinguish between creativity and innovation. Although creativity is a fundamental input for innovation, sometimes described as "the raw material" of innovation, it does not coincide with the latter. In fact, innovation requires not only a new idea regarding a product or a process, but also successful commercialization of the latter, so that consumers can pay a price for it. To conclude, we analysed the literature on the

relationship between trust and innovation, coming to the result that trust is generally conceived to be positively correlated with innovation, but this effect is often nonlinear and excessive amount of trust can even have a negative impact on innovation.

Our dissertation aims at following the literature strand regarding the relationship between trust and innovation between members of a dyad. We decided to study this phenomenon using tools from behavioral economics. We believe that trust and innovation should be studied in a context where the hypothesis of full rationality is relaxed and individuals are motivated by factors other than profit-seeking, even if we still consider the latter as a pivotal motive which cannot be ignored. We wish to study the psychological motives that lead high-trust individuals, societies and organisations, as well as organisations working in a high-trust environment, to possibly have a different attitude towards innovation.

Our work follows March's exploration-exploitation model (1991), thus it envisions innovation as a coordination problem, whereby members of a team or different organisations are called to come together and decide whether to continue with established organisational habits or try to follow a new path, which is riskier, but potentially more rewarding. In fact, although we acknowledge that some innovation may still originate from Renaissance-style masterminds, a wide body of literature (e.g., Powell, Koput and Smith-Doerr, 1996) has shown that in modern times innovation mostly stems from either large firms or networks of firms, and these organisations almost always have teams dedicated to innovation and R&D. Therefore, we consider it important to study how members of these teams and networks interact with each other and how the decision to innovate prevails over the decision to stick to a routine. Trust is expected to play a role in these interactions, which can be, for example, facilitating cooperation, increasing openness to new ideas even if coming from new or low-ranking employees, reducing the need of monitoring and all kinds of conflict, even including constructive conflict.

The most important gap that we observed in the literature is that it mostly relies on surveys or observational studies. Trust is often elicited with self-reported measures, obtained through questionnaires such as the World Values Survey's Rosenberg question, that is, "*Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?*" (Murtin et al, 2018, page 20). However, these measures suffer from two potential flaws. In fact, they are subject to the *social desirability bias,* that is, respondents tend to overstate their level of trust to please the experimenter. Moreover, the measure fails to capture the underlying motives that drive the decision to trust another person, such as positive expectations about the partner's trustworthiness or unconditional other-regarding preferences. Similar conclusions may be drawn for innovation, which is mostly elicited through either self-reported or result-based measures. An example of the latter is the number of patent applications. Instead, we decide to study innovation at its initial stage, when people decide whether to stick to an existing habit or attempt something new, which could lead to an innovation. The main novelty of our dissertation is that, in our second and third chapter, we decided to use experimental techniques to study the relationship between trust and innovation. Experiments provide many advantages. For example, the random assignment of subjects to different treatments helps reduce selection bias, the controlled environment of the experiment facilitates the isolation of the effect of the main explanatory variable, replication by other researchers is easier, and so is the identification of variables that are difficult to observe on the field, such as the psychological motives driving the choice to trust and

innovate. Trust is measured through the amount transferred by subjects in the trust game (Berg, Dickhaut and McCabe, 1995), while willingness to innovate is measured through subjects' choice between repeating a known task or attempting a new task, which involves creativity and is riskier and potentially more rewarding. Tools from game theory are used to simulate the interaction between team members, while tasks are meant to simulate real-life situations subjects might be involved in. In the third chapter, we also include information on the partner's level of trust. The goal is to control whether information on the partner's trust can act as a proxy for trustworthiness or expected reciprocity.

Although trust, innovation and the relationship between these two variables are well-established topics in the literature, we wish to provide some suggestions for future research. First, all these topics are highly interdisciplinary in nature, requiring tools from economics, management, psychology, political science and even chemistry and biology. Even if some satisfactory degree of collaboration between these different fields has already been reached, we advocate for an even greater level of interdisciplinarity in future research. Secondly, we believe that the so called "dark side of trust" should be further investigated, aiming to provide a more blurred but also more complex image of the relationship between trust and innovation. In fact, it is important for both entrepreneurs and policymakers to know whether the effect of trust on innovation is linear or whether there are some features of trust that have a different effect on innovation. For example, knowing that trust in others can lead to lower monitoring and constructive task conflicts can lead organisations to adopt countermeasures to help mitigate possible negative consequences on innovation and profitability. Similarly, it is of paramount importance to control for possible reverse causality. In fact, it might be the case that societies with better innovation performance also show higher levels of interpersonal and institutional trust due to the social development that innovation brings about. To conclude, as we stated earlier, we firmly believe that the psychological motives behind the choice of trusting and pursuing innovation deserve to be thoroughly investigated, and behavioral and experimental economics can provide useful tools for this important strand of research.

2 Trust and Innovation: An Experimental Study

2.1 Introduction

The aim of this study is to examine how social trust influences innovation. Specifically, we designed a two-stage online experiment in which participants engage in a strategic interaction within dyads. We wish to investigate whether individuals with higher levels of trust are more likely to attempt to innovate when successful innovation depends on coordination with their partner in a strategic interaction.

As we explained in the first chapter, the importance of trust and innovation in social and economic studies has grown significantly in the last decades. Both are multifaceted concepts which have a relevant impact on social and economic development. Trust in others is defined as "*a psychological state comprising the intention to accept vulnerability based upon the positive expectations of the intention or behaviour of another*." (Rousseau et al., 1998). This definition implies that trust has two building blocks, that is, risk and positive expectations on the behavior of others. Moreover, trust is affected by ambiguity about the probability of the counterpart's trustworthiness and strategic uncertainty about the actions and type of the partner. Trust has been defined "the first task of the ego" (Erikson, 1950) and it is widely recognised as of one the main determinants of economic growth (Knack and Keefer, 1997; Zak and Knack, 2001). Innovation is "*the implementation of a new or significantly improved product, good or service, or process, a new marketing method, or a new organisational method in business practices, workplace organisation or external relations*" (OECD, 2005), or, in other words, "invention plus exploitation". Innovation requires a creative idea that can be translated into something for which consumers can pay a price (Neto et al., 2019). Like trust, innovation is also based on the pillars of risk, uncertainty, and positive expectations. This happens because innovating always implies a greater risk of failure and uncertainty about the outcome than following a routine, but it also provides the possibility of higher returns. In the 21st century's "knowledge economy" (Drucker, 1969), that is, an economic system based on intellectual capital, the role of innovation is pivotal for growth and well-being (Schumpeter, 1934; Solow, 1956; Romer; 1986; Aghion and Howitt, 1992).

We investigated the relationship between trust and innovation by developing an online experiment run on the platform Prolific. This constitutes an element of novelty with respect to most existing studies. Experiments offer numerous advantages. For example, they allow researchers to better isolate the causal effect, and they are less affected by selection bias. Moreover, experiments are less susceptible to the social desirability bias and are easier to replicate. The experiment consists of two stages. In the first stage, we elicited participants' trust in others and risk attitude through two experimental tasks, that is, the trust game and the Bomb Risk Elicitation Task (BRET). In the second stage, participants played two rounds of a puzzle task, where they were required to reproduce a target figure. Through the repetition of the task, we aimed to simulate a habit, that is, a situation where participants learn to master easily through repetition. Later, we asked each participant to choose between an additional round of the puzzle task or a new task, which involved creativity and was riskier, but potentially rewarding. This binary choice was used as a measure of willingness to innovate. In this final phase of the experiment, participants were assigned to either the Matching or the Individual treatment. In the former, participants were matched with another player, and their payoff changed with the partner's choice. In fact, a bonus or malus could be added to a

participant's payoff according to her and the partner's choice, configuring a prisoner's dilemma situation. While cooperating on the creative task yielded the highest joint payoff, unilaterally deviating and playing the puzzle task always led to a higher individual payoff, so that mutual defection was the only Nash equilibrium. We hypothesized that high social trust could help participants escape this bad equilibrium. In fact, high-trust individuals could be more motivated to cooperate on innovation as they expect that the partner will follow them on innovation instead of betraying them. Moreover, they could be more inclined to think that they will get their fair share of the benefits of innovation, and they might perceive innovation as a more "moral" choice. Trust could help participants overcome fears related to strategic interaction and focus more on the possible benefits rather than the risk, ambiguity and uncertainty related to it.

Based on the literature, we formulated three hypotheses. First, participants in the Matching treatment could be less willing to innovate, as fear, ambiguity and strategic uncertainty could lead them to choose a safer option in a prisoner's dilemma setting. Secondly, the effect of trust on willingness to innovate could be nonsignificant when considering the whole sample, but positive and significant in the Matching treatment. In fact, only in a framework involving a strategic interaction trust can counterbalance the fears and biases that would otherwise trap dyads in a suboptimal equilibrium. Thirdly, the effect of trust on innovation in the Matching treatment could be nonlinear, that is, positive until a certain threshold of trust, then weaker or even negative. In fact, excessive levels of trust could lead participants to be more "conservative" and prefer to continue with a well-known habit. The experiment, including the hypotheses, the design plan, the sampling plan, the variables and the analysis plan, was preregistered on the OSF platform before the first experimental session.

Results only partially confirm our hypotheses. We prove that assignment to the Matching treatment significantly reduces participants' likelihood to choose the creative task. This is possible due to fears related to an impending strategic interaction, ambiguity aversion, strategic uncertainty, or betrayal aversion. While it is true that the effect of trust on willingness to innovate is nonsignificant while considering the whole sample, we do not find a significant effect of trust even in the Matching treatment. The effect of trust, albeit positive, is not strong enough to counterbalance the negative impact of assignment to the Matching treatment. However, we observe that participants do decide strategically, as we find a positive and significant correlation between expectation that the partner chose the creative task and willingness to innovate. Trust still does not seem to play a role in shaping participants' expectations about the partner's behavior. To conclude, we do not find evidence of a nonlinear relationship between trust and willingness to innovate in the Matching treatment, while we find a U-shaped relationship in the Individual treatment.

This paper is structured as follows. In the first section we summarized the literature on trust and innovation. In the second section we defined and explained our research hypotheses. In the third section we described the experimental design in detail, justifying it with references to the literature. In the fourth section we presented the results. The fifth section contains a discussion of the results and the limitations of our study, together with suggestions for future research. The last section was dedicated to conclusive remarks.

2.2 Literature review and contribution

In this section we will review the literature on trust and innovation, summarizing the trends described in the first chapter and positioning our study in the literature.

Trust and innovation are important concepts in economics and social sciences. From our literature review, it is clear that investigating these topics requires collaboration between different disciplines, including management, microeconomics, macroeconomics, behavioral economics, sociology and even chemistry or biology.

Researchers have found evidence that people do trust others even in strategic settings, e.g., in the trust game. Behavioral economics provides explanations for this choice. For example, people might decide to sacrifice part of their endowment due to other-regarding social preferences, including altruism and inequality-aversion (Evans and Krueger, 2009), social norms (Burnham, 2000; Hermann and Ariely, 2006), self-similar reasoning (Li, Turmunkh and Wakker, 2019), emotional factors (Dunning, Fetchenhauer and Schlösser, 2012; Di Bartolomeo and Papa, 2016b), or chemical and biological factors (McCabe et al., 2001; Rilling et al., 2001; Kosfeld et al., 2005; Zak, 2006). Trust has been shown to be positively correlated with economic development and growth (Knack and Keefer, 1997; Zak and Knack, 2001), investments (Zak and Knack, 2001), inclusive growth (Asongu and Gupta, 2015), competitive advantage for firms (Hunt, 2012), improved political outcomes (Knack, 2002; Bjørnskov, 2009) and the reduction of corruption (Rothstein and Uslaner, 2005; Morris and Klesner, 2010).

Behavioral economics explains innovation applying two distinct paradigms. The first is *choice under novelty,* whereby agents choose with no prior information and must avoid biases that can lead them away from innovation (Morrison and Pots, 2008). The second is *neuroinnovation,* whereby innovation is defined as a transformation process inside the brain, which translates unconscious stimuli into new products and services. The importance of innovation for economic development has been stressed by many authors since the beginning of the 20th century. Schumpeter's (1942) "creative destruction" is defined as "the essential fact about capitalism" (in Aghion and Howitt, 1992, p.324), while Solow (1956) and Romer (1986) find that technological progress and knowledge are the sole engines of long-term economic growth. The relationship between innovation, especially high-quality patents granted in the US, and economic growth is confirmed by Hasan and Tucci (2010) and Pradhan et al. (2016). In Aghion and Howitt (1992)'s model of growth through creative destruction the expected growth rate depends on the amount of research in the economy. However, firms might be disincentivized to invest in research if they fear that their monopoly power might be jeopardised by new innovations.

Having clarified the importance of trust and innovation in the social sciences literature, in the first chapter we focused on works investigating the relationship between these two variables. The literature on this topic is wide and multidisciplinary, and there seems to be a consensus on a positive relationship between trust and innovation. However, some authors find that excessive levels of trust can have a weaker or even negative effect on innovation. In the first chapters we classified the literature into four strands. In this section we are going to summarize the main findings, which we discussed in detail in the first chapter.

The first strand examines the effect of trust and social capital on innovation at the societal level. Social trust favours innovation outcome through many different channels. For example, it can facilitate knowledge sharing (Audretsch, 2019; Thompson, 2018; Lyu, Ma and Zhang, 2022), prevent egoistic behavior due to the fear of reputational loss (Akçomak and ter Weel, 2009; Lyu, Ma and Zhang, 2022), mitigate the agency problem in firms (Lyu, Ma and Zhang, 2022), help firm build a reputation for trustworthiness and reduce monitoring costs (Akçomak and ter Weel, 2009). Trust in other is correlated with the number of patents and the level of R&D expenditures, while institutional trust fosters high-tech exports (Dakhli and De Clercq, 2004). To conclude, Leon, Etsami and Nagi (2022) use a game-theoretic approach to prove that social groups characterized by higher social trust are more likely to adopt innovative over traditional products.

The second strand investigates how trust between members improves a team's ability to innovate. Clegg et al. (2002) find that trust that one's ideas will be listened to favors *idea implementation,* while trust that the organization will share the benefits of innovation predicts *idea suggestion*. Trustworthiness plays an important role in the formation of social network ties for idea generation and realization (Shazi, Gillespie and Steen, 2014). Creating a trusting environment within the firm (Lazány, 2017) and between members of the Top Management Team (Maccurtin et al., 2008) improves a team's innovation outcomes, as people who trust each other are better at cooperating, sharing knowledge and managing conflicts. However, although trust has a beneficial effect on innovation, it can also lead to a dangerous reduction in monitoring (Langfred, 2004) and in task conflict, that is, constructive conflict about how to conduct a specific task (De Clercq et al., 2009; Bidault and Castello, 2010). The consequence is that excessive levels of trust can have a weaker or even negative impact on innovation.

The third strand investigates the effect of trust on innovation when the latter involves cooperation between different organisations. Evidence suggests that social trust supports open innovation, acting as a substitute for formal contracts and reducing transaction costs and unintended knowledge spillovers (Brockman et al., 2018). Trust promotes the development of innovation ecosystems (Steinbruch et al., 2021) and it eases the exchange of ideas and knowledge (Teramoto and Jurčys, 2017). As regards ties between firms, social capital provides firms with an advantage in the formation of alliances for innovation (Gulati, 1998) and embeddedness, that is, a logic of exchange centred around social ties, increases a firm's performance (Uzzi, 1996). However, over-reliance on ties based on high trust can lead to dangerous consequences, such as being constrained by the network's set of rules, being blinded to extra-network opportunities, taking unnecessary risks, and preferring incremental over radical innovation (Uzzi, 1996; Molina-Morales, 2011; Budunchi, 2013). This "dark side of trust" (Molina-Morales, 2011) can cause excessive levels of trust to have a weaker or even negative effect on innovation.

The fourth strand investigates whether acting in a trusting environment gives firms a competitive advantage in innovation performance. Evidence shows that high-trust organizations outperform low-trust organisations due to a better resource exchange between business units and colleagues, higher loyalty, and engagement (Tsai and Ghoshal, 1998; Watson Wyatt, 2002), an easier transfer and creation of knowledge (Sankowska, 2013). They also avoid paying unnecessary "taxes", such as bureaucracy to ensure control and hinder opportunistic behavior (Covey, 2006). Moreover, firms acting in environments with high social trust produce more trademark applications, especially high-quality patents, and invest

more in R&D with lower volatility (El Ghoul et al., 2023; Dindaroglu, 2023). Firms belonging to industrial districts around Valencia have higher trust scores, and even within the district firms with higher trust and social capital have better innovation outcomes (Molina-Morales and Martínez-Fernández, 2010).

Overall, the literature indicates a positive correlation between trust and innovation. However, there are also potential drawbacks to excessive trust. For example, it can disincentivise constructive task-based conflict, reduce monitoring to dangerously low level and lead firms to miss market opportunities.

Our work aims to study the relationship between trust and innovation adopting a behavioral approach. Behavioral economics departs from the assumption that individuals are fully rational and profit-maximisers. Instead, agents can be influenced by cognitive biases and adopt mental shortcuts known as heuristics. Moreover, people can have unconditional other-regarding preferences like altruism and inequality-aversion (Cox, 2004; Di Bartolomeo and Papa, 2016a) or first-order beliefs, either negative or positive, about other people's trustworthiness. We believe that trust in others could be one of the mechanisms explaining the choice to innovate when people belonging to a small group are asked to coordinate to choose between continuing an existing habit or attempting something new. As we will explain, this recalls March (1991)' s exploration-exploitation dilemma, whereby organisations decide whether to refine existing products or processes or search for new solutions. In this second chapter, we adopted a theoretical framework that originated from behavioral economics and we tested it using tools from experimental economics. Using experiments constitutes a feature of novelty with respect to the existing literature. As we explained in the discussion section of the first chapter, the advantages of experiments include easier replicability, a controlled environment which facilitates the identification of causal effects and a lower susceptibility to selection bias. We used experimental measures of trust and willingness to innovate, which can help overcome the risk of social desirability bias and capture participants' behavior in tasks that resemble real-life situations. By nesting the choice between the known and the new task in a strategic situation involving a prisoner's dilemma, we aim to connect the literature on cooperation in social dilemmas with the literature on trust and innovation. In fact, we believe that in most teams the choice between following a habit and attempting to innovate can be seen as a strategic situation in which people with different preferences, seniority and background must coordinate on which path to follow. In our second chapter, we want to study the role played by trust in favoring the choice to innovate in a simplified version of this situation. As we will see in the experimental design section, participants belonging to the Matching treatment were paired with a partner and were asked to decide whether to continue with a known task or attempt a new task. Incentives were such that coordinating on innovation guaranteed the maximum payoff for the dyad, but defecting and choosing the known task maximized individual payoff. Our main research question is whether high-trust subjects are more or less likely to innovate in this strategic situation, and whether the effect of trust on willingness to innovate is linear. We will proceed to describe the experiment in detail in the following sections.

2.3 Hypotheses

Having provided an overview of the literature, we will now proceed to outline the hypotheses of this study. The general hypothesis is that the effect of trust on innovation will be positive, even if probably not linear. In dyads involved in a strategic interaction, high-trust members

will be more willing to innovate by choosing a new task instead of continuing an established habit. This design is based on March (1991)'s *exploration-exploitation dilemma,* whereby organisations need to decide whether to exploit, that is, to improve and invest in existing technologies or processes, or to explore, that is, to search for new technologies, processes, products. Exploitation guarantees safer but, on average, lower payoffs. Exploration exposes firms to higher risks: if it succeeds, rewards will be bigger, but if it fails, losses will be bigger. Organizations tend to prefer exploitation over exploration, but this is dangerous for their long-term profitability. This preference may be due to many factors, including loss aversion, low turnover, fast adaptation of new employees to the "organisational code", that is, the set of rules and habits of the firm, a strict hierarchical structure and resistance to change. Only positive expectations about the success of exploration can drive firms away from a possible bad equilibrium. In our paper, we hypothesize that high-trust participants belonging to the Matching treatment, that is, participants that are matched with a partner and involved in a strategic interaction, will be significantly more likely than low-trust subjects in the same treatment to choose exploration over exploitation. In fact, the literature shows that trust can favor knowledge exchange and the dissemination, suggestion, and implementation of new ideas. Trusting people are likely to have positive first-order beliefs about the trustworthiness of the partner, that is, to think that their partner will follow them in innovation, and they might want to send a signal of their trustworthiness and show that they are following a social norm. Moreover, reciprocal trust reduces the risk and fear of opportunistic behavior and the possibility of relationship conflict (Jehn and Mannix, 2001; De Cremer, 1999) between members of a team.

Hypothesis 1: *Willingness to Innovate will be higher in the Individual treatment than in the Matching treatment.*

As explained in the Introduction, subjects assigned to the Individual treatment are not matched with a partner. In fact, being involved in a strategic interaction could encourage individuals to choose the safer option-that is, to defect and continue playing the well-known puzzle task. This might happen due to many reasons. First, individuals in the Matching treatment could be subject to emotions such as anger (e.g., Dunn and Schweitzer, 2005; Delle Foglie, Papa and Spagnolo, 2025) and fear (e.g., Neto, Felipe and Bento Caleiro, 2019) which could drive them away from cooperating on innovation. Moreover, ambiguity aversion caused by the lack of information about the probability of cooperation of the counterpart and the nature of the new task (Li, Turmunkh and Wakker, 2019), strategic uncertainty regarding the behavior and type of the partner (Bruttel et al., 2023) and betrayal aversion (Delle Foglie, Papa and Spagnolo, 2025) might have led subjects to choose the safest option, that is, the puzzle task. Biases such as loss aversion, status-quo bias, and extremeness aversion (Potts, 2010) could have exacerbated by the strategic nature of the choice between the known and the unknown task. In fact, the likelihood of a loss for subjects who chose the new task does not only depend on the participant's bad performance or the difficulty of the task, but also on a possible opportunistic "deviation" by the partner, who can gain more if she decides not to follow the participant's attempt to innovate. Another issue is caused by the relatively more complex structure of the payoff matrix in the Matching treatment. As we will see later, the only scenario which does not include any bonus or malus with respect to the Individual treatment is the one in which both team members choose the puzzle task. In the Individual treatment, instead, the choice to innovate or continue with the puzzle task seems to

depend only on one's risk attitude, attitude towards change and self-perceived creativity, but no psychological factor induced by the interaction with a partner comes into play. Therefore, we expect willingness to innovate to be higher in the Individual treatment than in the Matching treatment. To verify this hypothesis, we ran three different tests. First, we compared the absolute frequency of the choice of the new task and the puzzle task in the Matching and the Individual Treatment. Second, we ran a Chi-squared test for independence to verify whether the difference in frequency was significant. Third, we ran a logistic regression whose dependent variable was Willingness to Innovate, that is, a dummy equal to 1 if the subject chose the new task, and the regressor was Matching, a dummy equal to 1 if the subject was assigned to the Matching treatment.

Hypothesis 2: *Considering the whole sample, the effect of Trust on Willingness to Innovate will be non-significant, while in the Matching treatment the effect of Trust will be positive and significant.*

The effect of trust in others on innovation should be visible only in cases where an interaction is present. In fact, we defined trust as the intention to accept vulnerability based on positive expectations on the behavior of a partner. Interactions always involve a certain degree of vulnerability to opportunistic behavior and betrayal by the partner. In our experiment, this would mean deviating from cooperation by choosing not to innovate when the other party has chosen to do so. In fact, in the Matching treatment the dyad maximizes its joint payoff by coordinating on innovation, but if one member decides to deviate and play the puzzle task, she will obtain the maximum gain, while the partner who does not deviate will suffer the maximum loss. We thus expect the coefficient of trust to be non-significant when considering the whole sample, which includes subjects that are not matched with a partner. In the Matching treatment, instead, we hypothesize that high-trust subjects could be more likely to innovate than low-trust subjects. In fact, high-trust participants could expect cooperation from the partner and believe that they will get a share of the benefits of innovation, that is, the bonus added to the payoff of the creative task (Clegg et al., 2022). They might want to signal their trustworthiness by deciding to cooperate on innovation, with the goal of either inducing trustworthy behavior by the counterpart or showing that they adhere to social norms (Evans and Krueger, 2009). High-trust participants could in fact see the prisoner's dilemma in terms of morality rather than might (Van Lange and Kuhlmann, 1994) and perceive the choice to trust as moral. They could possibly be led by self-similar reasoning and the false consensus effect to believe that other people are also trusting and trustworthy (e.g., Li, Turmunkh and Wakker, 2019) and thus decide to cooperate on innovation. If this happens, guilt aversion could also motivate high-trust participants to choose to innovate to avoid letting the partner down. In an actual team, people with higher trust could be more willing to express their innovative ideas. Trust in others could help overcome fear of ridicule (Neto, Felipe and Bento Caleiro, 2019) and betrayal aversion (Delle Foglie, Spagnolo and Papa, 2025) and lead participants towards innovation. De Cremer (1999) shows that trust could help reduce fear of being exploited in a social dilemma, because high level of trust might be associated with the perception of fairness and collective efficacy, that is, subjects would believe that others are fair and "they can solve their problems and improve their lives through concerned effort" (Bandura, 1986; in De Cremer, 1999, p. 156).

To test this hypothesis, we ran a series of logistic regressions whose dependent variable was Willingness to Innovate and whose regressors were Trust and an interaction term between

trust and the dummy Matching. If the coefficient of the interaction term is positive, Hypothesis 2 will be confirmed:

$$Willingness\ to\ Innovate_i = \alpha + \beta Trust_i + \gamma Matching_i + \delta Trust_i * Matching_i + \varepsilon_i$$

Hypothesis 3: *The effect of Trust on Willingness to Innovate in the Matching treatment could be nonlinear: after a certain threshold, the effect of Trust on Willingness to Innovate might become weaker or even negative.*

The literature provides some evidence that the effect of Trust on Innovation might not be linear. This could be attributed to many factors. For example, an excessive level of trust could lead to the suppression of *task-based conflict* (Bidault and Castello, 2010; De Clercq et al., 2009). Subjects would become accommodating towards their partners and unwilling to confront them when they have different views about the task. Moreover, when trust between team members is too high, monitoring could be reduced to dangerously low levels (Langfred, 2004). This could lead teams to lose potential market opportunities (Uzzi, 1996) and stifle radical innovation (Budunchi, 2014). In our experiment, we hypothesize that subjects with a high level of social trust will be more likely to choose the creative task when matched with a partner, but after a certain threshold, this effect might be weaker or even negative. In fact, individuals who are excessively trustful may prefer to "exploit" and play the puzzle task, as they expect their partner to do the same. Moreover, it could also be the case that people with extremely high trust are generally more wary of innovation, especially radical innovation, and more risk or ambiguity averse. To test this hypothesis, we ran a logistic quadratic regression whose dependent variable was Willingness to Innovate and whose explanatory variables were the interaction terms Trust*Individual and Trust*Matching and the same interaction terms squared. We decided to study the nonlinearity of trust separately in the two treatments; therefore, we did not include a simple Trust regressor. The regression was the following:

$$\begin{aligned} Willingness\ to\ Innovate_i = {} & \alpha + \beta Trust_i * Individual_i + \gamma Trust^2{}_i * Individual_i + \\ & \delta Matching_i + \zeta Trust_i * Matching_i + \eta Trust_i^2 * Matching_i + \\ & \varepsilon_i \end{aligned}$$

If η is negative, the curve is concave and the effect of trust on willingness to innovate is positive until a certain threshold, then negative, thus confirming our hypothesis.

2.4 Experimental design

This experiment involved 208 participants and was conducted online through the platform Prolific. Compared to laboratory experiments, Prolific allows us to obtain quicker and reliable data and to reach a higher number of participants, without being limited to university students. This comes at the cost of a lower level of control on participants with respect to laboratory experiments, but we believe that the benefit of a larger and more varied subject pool justifies this potential issue. Given that the creative task requires an excellent level of English language proficiency, we decided to recruit participants from Great Britain. A pilot has also been undertaken on Prolific. The experiment was generated on oTree, an open-source platform for web-based interactive tasks. Players were informed that their anonymity would have been guaranteed throughout the whole experiment.

The experiment consisted of two stages, undertaken with a time distance of four days. In the first stage participants were paid a show-up fee of £ 1, while in the second stage they were paid a show-up fee of £ 1.5. Bonus payments generated by each participant's performance in

the tasks were added. Only participants who showed up in both stages of the experiment received the full payment.

In the first day, we elicited participants' idiosyncratic preferences, that is, trust in others and risk attitude, through experimental procedures such as the trust game and the Bomb-Risk Elicitation Task (BRET). In the second day, we asked players to familiarize with a task and then offered them the opportunity to choose between playing another round of this task or attempting an unknown task, which was riskier, but more rewarding. Participants were either assigned to either the Individual or the Matching group.

At the end of the experiment, we collected data on the subjects' self-perceived creativity, age, gender and nationality. In each task, players gained a sum in Experimental Currency Units (ECU), which was then converted in pounds at the following rate: 1 ECU=$ 0.01. This is the structure of the experiment:

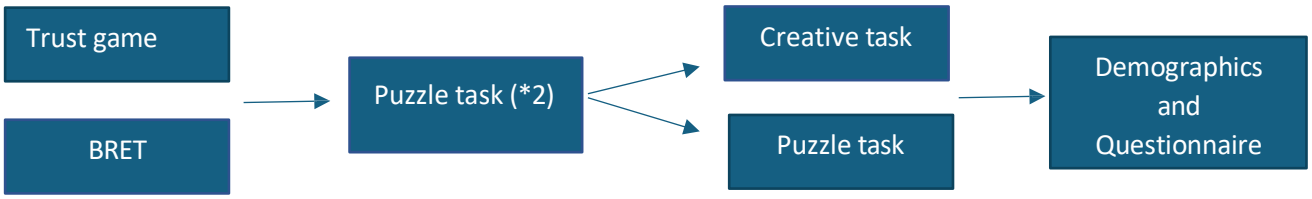


Figure 1: Structure of the experiment

2.4.1 First stage

In the first stage, we elicited participants' idiosyncratic features that may influence their willingness to innovate. We focused on trust and risk attitude, while self-perceived creativity was measured at the end of the experiment.

We measured each participant's trust in others through the trust game. We decided to use an experimental measure, because survey measures traditionally employed, such as the World Values Survey's *Rosenberg Question*, suffer from issues such as *social desirability bias,* that is, respondents give untruthful answers to please the experimenter. Instead, in the trust game participants are expected to act almost as they would do in everyday life, thus returning a more "honest" measure of trust in others. (Murtin et al., 2018).

The trust game, first developed by Berg, Dickhaut and McCabe (1995), is a sequential game played by two players, A and B. Both start with an endowment of 10 ECU. Player A is the first to move and decides how much of her endowment to transfer to player B. The amount transferred is then tripled, and B can decide how much of this sum to return to player A. It must be noted that the sum returned by B cannot be higher than the sum transferred by A multiplied by 3, so that if A does not transfer anything, B cannot return anything. Therefore, if we indicate with t the amount transferred by A and r the amount returned by B, A's final payoff is $10 - t + r$, B's final payoff is $10 + 3t - r$, with $r \leq 3t$. The amount transferred by A is a measure of trust, while the amount returned by B is a measure of trustworthiness. Although the Nash equilibrium of the trust game is reached when no player transfers anything to her partner, a wide body of literature has provided evidence of prosocial behavior (Evans and Krueger, 2009). Here is a figure describing the trust game. Players were shown this figure in the instructions of the task. The task was presented to players as "Task 1" to avoid priming.

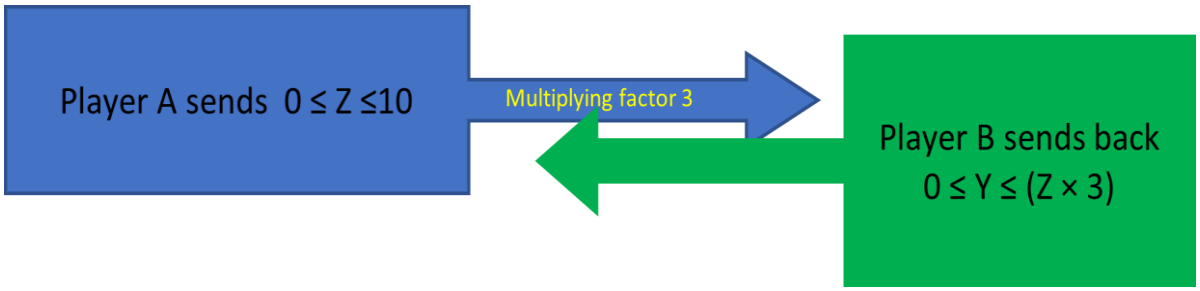


Figure 2: The trust game

In this version of the trust game, all participants played both as player A and player B. Therefore, they were asked at the same time how much of their endowment they would transfer to another player and how much they would return for each possible amount received by another player. Each player made eleven choices, one as player A and 10 as player B. Therefore, if t is the amount transferred as player A and r is the amount returned as player B, the choice vector of player i is the following:

$$(t_i, \overline{r_{i_{t_j}}}), \text{ where } \overline{r_i} = (r_{i_1}, r_{i_2}, r_{i_3}, r_{i_4}, r_{i_5}, r_{i_6}, r_{i_7}, r_{i_8}, r_{i_9}, r_{i_{10}}).$$

This procedure is called strategy method (Selten, 1967), and it is employed in many works, including Sapienza, Toldra-Simats and Zingales (2013) and Murtin et al. (2018). To calculate final payoffs, half of the players were randomly assigned to the role of player A, the other half were assigned to the role of player B, then pairs were randomly formed. For each pair (i, j), where i is player A and j is player B, the payoff vector was as such: $(10 - t_i + r_{j_{t_i}}, 10 + 3t_i - r_{j_{t_i}})$, where $r_{j_{t_i}}$ is the amount returned by player j for the amount t_i transferred by player i. When playing as player B, players were asked to fill in the third column of a table like the following, knowing that they could not return more than they received:

If Player A sent to you	**You received**	**You send to A**
1	3	
2	6	
3	9	
4	12	
5	15	
6	18	
7	21	
8	24	
9	27	
10	30	

Table 1: Choices as player B with the strategy method

Pairs formed for this stage were not the same as those formed for the second stage. Players were informed that they would play both as player A and player B and that their payoff would be determined through a random assignment of both role and partner during the instructions. After the instructions, players were asked comprehension questions regarding the number of choices they could make as player A and player B and the payoff they would gain for specific combinations of amounts sent by player A and returned by player B. Players could not proceed in the task if they did not answer the comprehension questions correctly. For the trust game, each player could earn a maximum of 40 ECU, which corresponds to £ 0.4.

Secondly, we measured risk attitude using the Bomb Risk Elicitation Task (BRET). This task was first developed by Crosetto and Filippin (2013). Again, we decided to use an experimental measure rather than a questionnaire to avoid the risk of social desirability bias. In the BRET, participants were shown a screen with 100 boxes. They were told that each of them contained 1 ECU, except for one, which contained a bomb. Participants had to choose how many boxes to collect. Their earnings increased linearly with the number of boxes collected, but they were equal to 0 if they collected the box with the bomb. The task was presented to players as "Task 2" to avoid priming. Below is the screen that participants saw. They needed to write the number of boxes they want to collect, then click on "Stop" and "Continue". As shown by Crosetto and Filippin (2003), the BRET has many advantages compared to similar measures of risk attitude. From a practical perspective, the BRET allows for a precise estimation of both risk aversion and risk seeking, combined with a simple and intuitive design. From a theoretical point of view, the BRET has two main advantages. First, loss aversion is not a confounding factor, because the task is entirely defined in the gain domain. Secondly, the clear distinction between risk aversion and loss aversion allows us to clarify gender differences in risk attitudes. In fact, females do not show a different risk aversion, instead they have a stronger loss aversion.

In the first stage, players were paid according to their performance in either the BRET or the trust game, for a maximum of £ 0.99. Players were informed that they would have known their payoff and payment only at the end of the experiment.

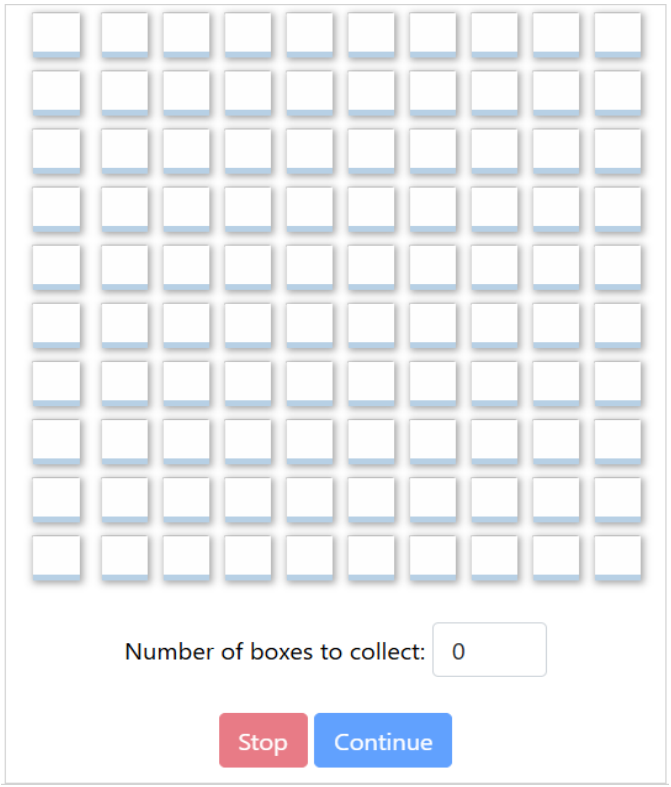


Figure 3: The BRET

2.4.2 Second stage

2.4.2.1 Overview

In the second stage, players were involved in a series of real-effort tasks, that is tasks where "*participants must actually work, be it manual, mental, or both to some degree, to achieve outcomes*" (Carpenter and Huet-Vaughn, 2019). We chose to use real effort tasks because they allow us to reproduce real-life feelings, like boredom, fatigue, anxiety, and excitement, which are usually excluded from traditional abstract experiments.

The first stage was divided into a training phase and an active phase. In the training phase, players familiarised with a puzzle task by repeating it for two rounds. In the active phase, they chose whether to play a third round of the Puzzle Task or attempt an unknown creative task, which could yield higher returns with higher risk. Further details will be provided in the following subsections. The puzzle task and the new task were presented to players as respectively Task 3 and Task 4 to avoid priming. As explained before, the structure of the second stage recalls the Exploration-Exploitation dilemma, with the puzzle task representing exploitation and the creative task representing exploration (March, 1991). Both the puzzle task and the creative task were played alone. Before each task, participants were asked some comprehension questions and could proceed in the experiment only if they answered correctly.

In this module players were randomly assigned to two treatments. In the Matching treatment, after the choice between the puzzle and the creative task, players were paired with another participant and their payoff partially depended on their partner's choice (but not on the partner's performance). Pairs were not the same as in the trust game. In the Individual treatment, instead, players simply chose between the puzzle and the creative task, without

being matched to anyone. This was intended to rule out the strategic interaction component from participants' choices.

Individual	**Matching**
No pairs	Subject paired randomly with a partner
Tasks played alone	Tasks played alone
No strategic interaction	Strategic interaction: subject's payoff depends on partner's choice

Table 2: Summary of Individual and Matching treatments

2.4.2.2 The Puzzle Task

In the training phase, each participant completed two rounds of a puzzle task, one of which was randomly selected for final payment. In each round, participants saw a screen with two grids; the right-hand grid displayed a target figure composed of black and white blocks, called "pixels", while the left-hand grid initially appeared blank, consisting only of white pixels. The figure in the right-hand grid constituted the target figure, which participants needed to reproduce in the left-hand grid. To complete the task, subjects were instructed to click on white pixels in the left-hand grid to turn them black. The number of moves was unlimited; however, each round had a time limit of 40 seconds. Because the target figures were small and relatively simple to reproduce, a round was considered successful only if the subject made no mistakes. In case of success, she would earn 180, while in case of failure she would earn 100 ECU. At the end of the experiment, one round of the Puzzle Task was randomly selected for payment, with a conversion rate of 1 ECU= £ 0.01.

An example of the task is provided in the figure below:

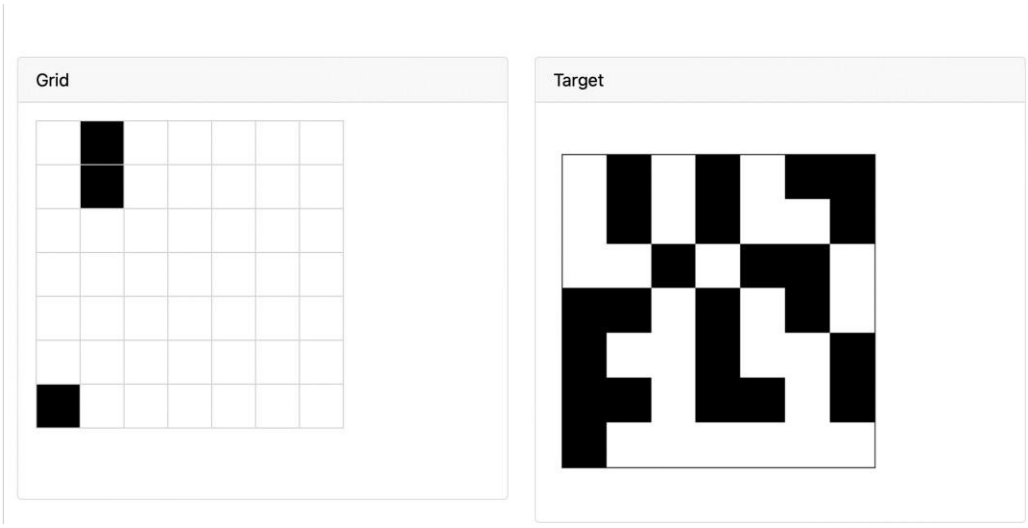


Figure 4: The Puzzle Task

The repetition of the task for two rounds should allow for a refinement of choices through learning by doing, so that a routine is formed. This is in line with the process of brain learning known as "repetition-suppression", whereby the more the brain receives similar information, the less effort it will make to interpret it. Consequently, after a certain number of repetitions most of the work will be performed by only a fraction of the network of neurons. (Neto et al., 2019). This is needed to recreate the "lock-in" effect that usually hinders innovation, as people

stick to the habit that they feel good at performing due to a path dependency effect (Foxon, 2014).

2.4.2.3 Choice

After the second round of the puzzle task, the active phase began. Each player was required to choose whether to play another round of the puzzle task, called Task 3, or to play a round of a new task, called Task 4. Participants were informed that if they chose the new task, they would have the opportunity to earn more than in Task 3, but they would earn less in case of failure. Moreover, they were told that being creative could improve the probability of success in the new task. As we will explain in the following subsections, the new task, called *Divergent Association Task* (DAT) yielded a payoff of 200 in case of success and 80 in case of failure, while playing another round of the puzzle task still yielded a payoff of 180 in case of success and 100 in case of success. The payoff structure of the new task was more extreme than the puzzle task, thus being riskier, but potentially more rewarding. This is in line with March (2019)'s exploration-exploitation dilemma.

Willingness to Innovate was measured as a dummy variable equal to 1 if the participant decided to attempt the new task, 0 otherwise. Participants were randomly assigned to either the Individual or the Matching treatment. Subjects assigned to the Individual treatment were directed to the choice in the following page of the experiment. Instead, subjects assigned to the Matching treatment were told that they would be matched with another partner, who would not be the same partner as in the trust game, and her payoff would also depend on the choices of the partner. In the following page, subjects assigned to the Matching treatment were instructed about the bonus-malus system associated with the simultaneous choice with the partner, which we will describe in the next subsection. However, participants in the Matching treatments played the chosen task alone, without communicating with the partner. Therefore, in the Matching treatment participants were subject to a triple strategic uncertainty, regarding the behavior of the partner, her identity and type and the nature of the new task.

2.4.2.4 Matching treatment

As anticipated before, in the Matching treatment each participant was randomly paired with another participant and her payoff depended not only on her performance, but also on her partner's choice. Pairs were not the same as those defined for the trust game. In case the number of participants assigned to the matching group was odd, one participant would have no partner. To solve the issue, her payoff would have been calculated by matching her with one of the participants who have already been assigned to a group, so that her payoff depended on the partner's choice, but the opposite was not true.

We decided to create pairs to simulate an exploration-exploitation dilemma happening in a team, where the decision requires a coordinated and collective choice. Each subject had to consider whether to "innovate", that is, to try the creative task, based on their expectation on whether their partner will follow them in innovation. We decided to study willingness to innovate in dyads because evidence shows that in contemporary societies innovation mostly stems from collective decision-making happening in large firms or networks of firms practicing open innovation (e.g., Powell, Koput and Smith-Doerr, 1996), rather than from single bright individuals. Innovating in these teams requires coordination between members, and trust could play a role in facilitating the choice to innovate.

In the Matching treatment, a bonus or malus could be added to each participant's performance according to her partner's choice. Please note that, for the sake of simplicity, the partner's performance in the chosen task did not influence participants' payoff. The bonus-malus system was defined as such:

- If both players decided to attempt the DAT, they gained the payoff of the DAT plus a bonus of **40**. The bonus is a reward for coordination. In fact, by sharing ideas, effort, and money it is easier to innovate.
- If one player decided to attempt the DAT and the partner decided to stick to the Puzzle Task, the former got a malus of **40**, while the latter got a bonus of **100.** In fact, deciding to innovate alone means to sustain all the costs of innovation with a lower probability of success, while leaving the partner alone to innovate means to free-ride and have a safer payoff with a known task, that is, the puzzle task.
- If both players chose the puzzle task, no bonus nor malus was added.

Participants in the Matching treatment were informed of this bonus-malus system, but not of the motivation laying underneath it. They knew what bonus or malus they and their partner would gain in all combinations of choices, and that their payoff in the chosen task depended partially on the partner's choice, but not on the partner's performance. A comprehension question was asked before subjects could make their choice, to ensure they understood the bonus-malus system. Moreover, after playing the chosen task, each participant was asked what task she thought the partner had chosen, and whether she thought the partner had succeeded in the task. We thus obtain the dummy variables *Expectations about partner's choice* and *Expectations about partner's success.*

More formally, the payoff structure is the following, where $\pi_1(ct)$ and $\pi_1(pt)$ represent the payoff of player 1 in respectively the creative and the puzzle task and $\pi_2(ct)$ and $\pi_2(pt)$ represent the corresponding payoffs of player 2. Below is the resulting payoff table. For the sake of simplicity, participants were not shown the payoff table, but only the description of the bonus-malus systems.

	CT	**PT**
CT	$\pi_1(ct) + 40;$ $\pi_2(ct) + 40$	$\pi_1(ct) - 40;$ $\pi_2(pt) + 100$
PT	$\pi_1\,(pt) + 100;$ $\pi_2(\mathrm{ct}) - 40$	$\pi_1(pt); \pi_2(pt)$

However, it must be noted that each player i has a probability p_i of succeeding in the creative task and q_i of succeeding in the puzzle task. Therefore, $\pi_i(ct) = p_i * 200 + (1 - p_i) * 80$ and $\pi_i(pt) = q_i * 180 + (1 - q_i) * 100$. Adding the possible bonus or malus, the potential payoff structure including participants' probabilities of failure and success:

	CT	PT
CT	$p_1 * 240 + (1 - p_1) * 120;$ $p_2 * 240 + (1 - p_2) * 120$	$p_1 * 160 + (1 - p_1) * 40;$ $q_2 * 280 + (1 - q_2) * 200$
PT	$q_1 * 280 + (1 - q_1) * 200;$ $p_2 * 160 + (1 - p_2) * 40$	$q_1 * 180 + (1 - q_1) * 100;$ $q_2 * 180 + (1 - q_2) * 100$

It can be proven that for each p_1, p_2, q_1, q_2 we always obtain a prisoner's dilemma. To avoid priming, players were not informed that they would play a prisoner's dilemma, instead they only got to know that they would play another individual task, and their payoff would partially depend on the partner's choice. In the prisoner's dilemma originated by the choice, the Nash equilibrium is reached when both players decide not to innovate. Playing the puzzle task is a dominant strategy, that is, it leads to a higher payoff notwithstanding the choice of the partner. However, coordinating on innovation leads to a Pareto-efficient outcome, whereby no player can earn more without harming the partner: if they succeed, they both earn 240 ECU. This pair of strategies is not an equilibrium, because each player can earn more by "defecting" and deciding to play the Puzzle Task, leaving the partner with a much lower payoff. We hypothesize that trust could help dyads escape the bad equilibrium and reach the Pareto-efficient outcome, as it could lead participants to believe that the partner would follow them in innovation.

2.4.2.5 Individual treatment

In the individual group, after the two rounds of the puzzle task, each player decided whether to play another round of the puzzle task or the DAT. However, they were not matched with another participant, therefore the element of trust was not present. In fact, the choice could depend on risk attitude, self-perceived creativity, and other factors, but trust in others should not have come into play. Success in the DAT yielded 200 ECU, failure yielded 80 ECU. Success in the puzzle task yielded 180 ECU, failure yielded 100 ECU. No bonus nor malus were added to the payoff obtained in the chosen task.

2.4.2.6 Divergent Association Task (DAT)

The Divergent Association Task (DAT), first developed by Olson et al. (2021), is a task aimed at measuring creativity. The task measures divergent thinking, which is the ability to generate multiple solutions for solving a problem.

In the DAT, each participant was asked to generate 10 nouns that needed to be as different from each other as possible in all meanings and uses of the words. An algorithm called GloVe (Pennington et al., 2014) calculated the semantic distance between each pair of words. Words that are used in similar contexts, like cat and dog, would have a small semantic distance, while words that are more distant from each other, like hippo and train, would have a higher semantic distance. Some rules apply regarding the choice of words:

1) Only single words can be used. Composite words like *cul de sac* are also accepted.
2) Only nouns can be used. Avoid adjectives, adverbs, verbs etc.

3) Proper nouns are not allowed.
4) Specialized vocabulary (technical terms) is not allowed.
5) Participants must think of the words on their own (e.g., without looking at objects in your surroundings).
6) The time constraint for the task is 4 minutes.

The DAT score for each participant was calculated by keeping the first seven valid words and calculating the semantic distance between all 21 possible pairs of the seven words, taking the average, and then multiplying it by 100. The full algorithm is available online (https://osf.io/vjazn/). Scores range from 0 to 200. However, scores rarely exceed 100 or are less than 50. The score could thus be interpreted as such: under 50 it is poor, between 75 and 80 it is average, and more than 95 is very high. In our experiment, we asked all players to provide 10 nouns and asked GloVe to calculate the score for each player. We then calculated the median score for our sample, and each player's performance was considered successful if her score was bigger or equal to the sample median. In case of success the participant gained 200 ECU, in case of failure she gained 80 ECU. The rate of conversion was 1 ECU= £ 0.01. Before playing the task, participants were asked some comprehension questions. Moreover, they were asked to rate their probability of success in the DAT from 1 to 10. The explanation of the DAT was only provided to subjects that chose it. An example of the task is shown below:

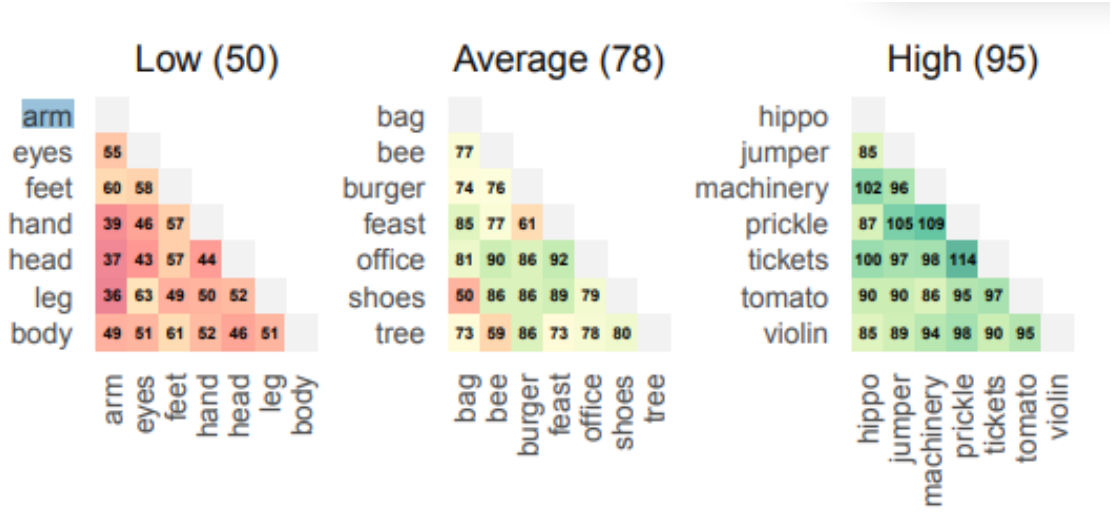


Fig. 1. Examples of participant responses and their corresponding DAT scores. The score is the transformed average of the semantic distances between each pair of words.

Figure 5: The DAT score

Olson et al. (2021) describes the main advantages of the DAT. First, it is automatic and objective. Second, it allows to collect large samples with little effort and no rater bias, that is, with no need of a judge to assesses each participant's level of creativity. Moreover, the scoring is absolute and not sample-dependent and avoids experimental and reliability biases by giving an open-ended prompt and using a model trained on a global website data. By requiring all participants to generate the same number of responses, we ensured that the score could not be inflated by fluency, that is, the number of uses a participant manages to generate from a certain tool. Moreover, the task is short and enjoyable.

2.4.3 Demographics and self-perceived creativity

After the second stage of the experiment, we measured each participant's self-perceived creativity. In fact, self-perceived creativity can be a confounding variable, influencing participants' willingness to innovate. It was measured through a questionnaire based on the creative potential subscale of the Creative Potential and Practiced Creativity scale (Mittone, Morreale and Vu, 2022). In this questionnaire, subjects were required to state how much they agreed with 6 statements, on a scale from 1 (strongly disagree) to 5 (strongly agree).

The questions were the following:

1. I feel that I am good at generating novel ideas.
2. I have confidence in my ability to solve problems creatively.
3. I have a knack for further developing the ideas of others.
4. I am good at finding creative ways to solve problems.
5. I have the talent and skills to do well in my future job.
6. I feel comfortable trying out new ideas.

To conclude, subjects were asked their age, gender, nationality and educational. At the end of the experiment, participants received a payment equal to the sum of:

- A total show-up fee of £ 2.5
- The payoff of either the trust game or the BRET
- One of the two rounds of the puzzle task, randomly selected
- The payoff of the chosen task, including the possible bonus or malus in the Matching treatment.

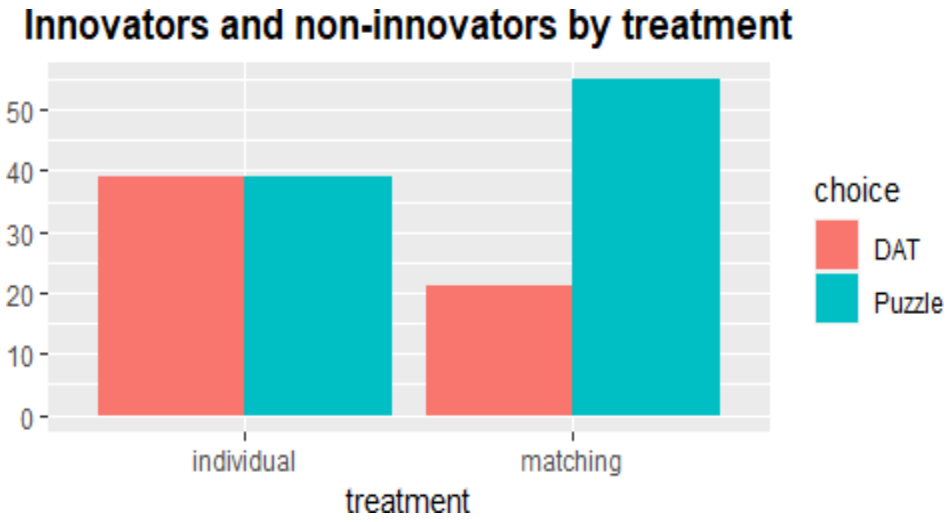
2.5 Results

For our data analysis we decided to run a series of regressions on R and to use instruments such as the Chi-squared test for independence. The main dependent variable, *Willingness to Innovate*, is a dummy equal to 1 if the participant chose the creative task (DAT), 0 if she chose the puzzle task. Therefore, we decided to employ a logistic regression to capture the effect of Trust on Willingness to Innovate. We also created a dummy equal to 1 if the participant was assigned to the Matching treatment, 0 if she was assigned to the Individual treatment, and an interaction term between *Trust,* that is, the amount sent in the trust game, and the aforementioned dummy variable. This variable was called *Trust*Matching.*

2.5.1 Hypothesis 1

We wished to investigate whether participants in the Individual and the Matching treatment chose the DAT with a different frequency. For this purpose, we ran three different tests, which led to the confirmation of Hypothesis 1.

First, we compared the absolute frequency of the choice of the DAT, that is, the creative task, and the puzzle task among the 154 participants who took part in the second stage. While in the Individual treatment participants were equally divided between the DAT and the puzzle task, in the Matching treatment 39 participants chose the puzzle task and only 21 participants chose the DAT. Here is the bar plot:



To check whether this difference was significant, we ran a Chi-squared test for independence. It is a nonparametric test to determine whether two categorical values are independent or related to each other. Given the contingency table showing the number of observations in each treatment, the null hypothesis suggests that the absolute frequency of the choice of the DAT in the Matching group should be equal to the absolute frequency of the choice of the DAT in the Individual treatment. Rejecting the null hypothesis would mean that assignment to the Matching treatment played a role in participants' choice. With O being the observed frequency and E being the expected frequency, the test statistic is the following:

$$X^2 = \sum \frac{(O - E)^2}{E}$$

As the p-value is smaller than 0.01, we can reject the null hypothesis and suggest that assignment to the Matching treatment played a significant role in participants' choice, dissuading them from innovating.

Chi-squared test for independence

X-squared	7.815***
Degrees of freedom	1
p-value	0.007

*Chi-squared test with Yates' continuity correction * p<0.1, ** p<0.05, *** p<0.01*

To further prove this hypothesis, we regressed Willingness to Innovate on the dummy *Matching*. Given that the dependent variable is a dummy, we ran a logistic regression. Results show a negative and significant effect of Matching on Willingness to Innovate, providing further evidence in favor of the hypothesis that people in the Matching treatment were less likely to innovate.

Willingness to Innovate across treatments

	Dependent variable:
	Willingness to Innovate
Matching	-0.963*** (0.342)
Constant	0.000 (0.226)
Observations	154
Log Likelihood	-98.863
Akaike Inf. Crit.	201.726

Logistic regression estimates.

*Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

2.5.2 Hypothesis 2

To control for Hypothesis 2, we ran three regressions. The first was a simple logistic regression whose dependent variable was Willingness to Innovate and whose explanatory variable was Trust:

$$Willingness\ to\ innovate_i = \alpha + \beta Trust_i + \varepsilon_i$$

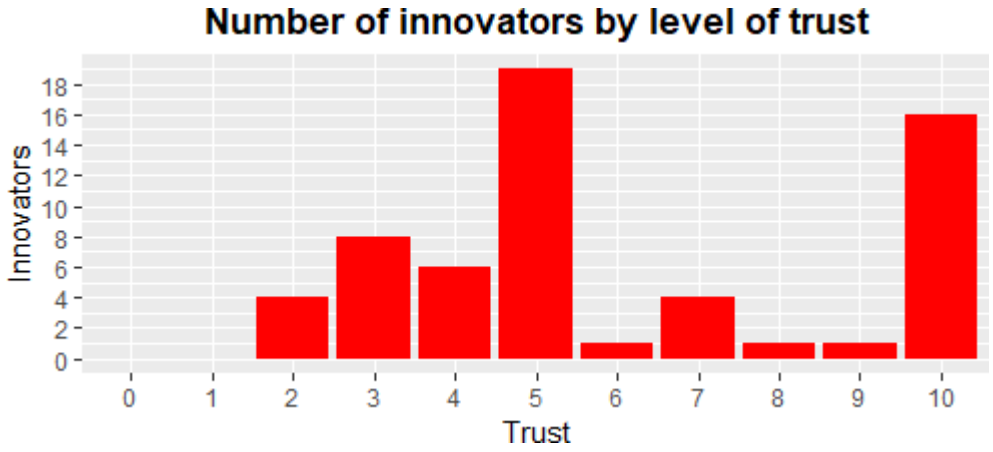
As expected, the coefficient of Trust is not significant, as we considered the whole sample, including the Individual treatment.

Trust and Willingness to Innovate

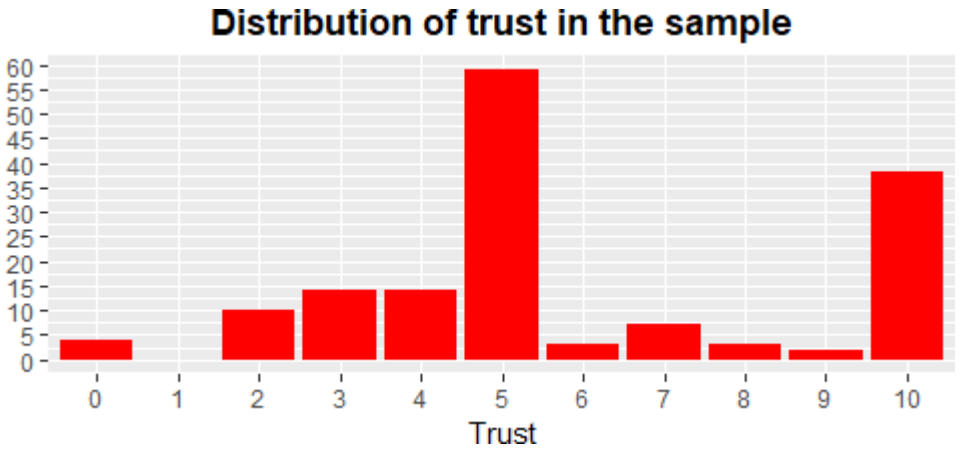
	Dependent variable:
	Willingness to Innovate
Trust	0.037 (0.059)
Constant	-0.669* (0.389)
Observations	154
Log Likelihood	-102.763
Akaike Inf. Crit.	209.525

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

We created a bar plot to represent the absolute frequency of participants who chose to innovate for each level of trust. No clear trend is visible, except for two spikes when trust was equal to 5 and 10.



The spikes correspond to those observed for the distribution of levels of trust across the sample, as shown here:



We proceeded by adding Matching as a control. Its coefficient remains negative and significant, thus confirming the negative effect of assignment to Matching on Willingness to Innovate.

Trust and Willingness to Innovate

	Dependent variable:
	Willingness to Innovate
Trust	0.032 (0.062)
Matching	-0.956*** (0.343)
Constant	-0.189 (0.433)

Observations	154
Log Likelihood	-98.732
Akaike Inf. Crit.	203.464

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

To capture the effect of Trust in the Matching group we added the interaction term *Trust*Matching* to our regression, together with the main effects Trust and Matching. We then proceeded to add controls, summarized in the matrix X_i, for a robustness check. Therefore, we ran the following regression:

$$Willingness\ to\ innovate_i = \alpha + \beta Trust_i + \gamma Trust_i * Matching_i + \delta Matching_i + \theta X_i + \varepsilon_i$$

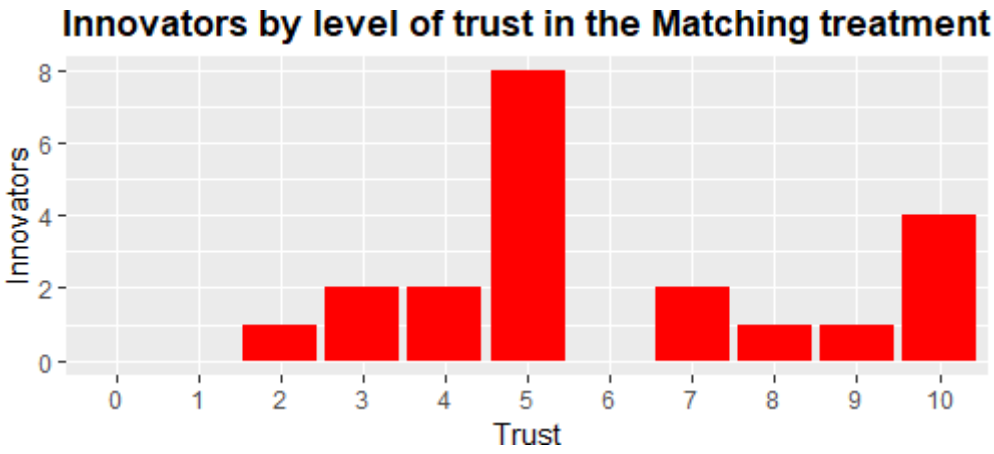
As can be seen from the tables below, we do not find evidence of a significant effect of the interaction variable on Willingness to Innovate. This means the effect of trust in the Matching treatment is not significant. However, one can observe that the coefficient of *Trust*Matching* is positive, while the coefficient of *Matching* is negative. Even though both coefficients are not significant, we could hypothesize that trust in others partially weakens the negative effect of assignment to Matching on Willingness to Innovate, although not enough to counterweight it.

Interaction and Willingness to Innovate

	Dependent variable:
	Willingness to Innovate
Trust	0.011 (0.087)
Matching	-1.197 (0.810)
Trust*Matching	0.041 (0.123)
Constant	-0.068 (0.568)
Observations	154
Log Likelihood	-98.678
Akaike Inf. Crit.	205.356

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

The bar plot showing the distribution of the number of innovators by level of trust in the Matching treatment does not reveal any trend, except for a spike when trust is equal to 5.



To perform another check, we introduced some controls. *Risk attitude* measures the number of boxes collected in the BRET, *Average trustworthiness* measures the average amount returned in the trust game, *Puzzle task ability* is the average payoff obtained in the first two rounds of the puzzle task. We also added *Gender* and *Educational Attainment*. Only Average Trustworthiness is negatively and significantly correlated with Willingness to Innovate. We did not include the variables *DAT score* and *Expected success in the DAT,* as they are only defined when *Willingness to Innovate* is equal to 1. We also did not include *Nationality*, as almost all participants were British.

Interaction and Willingness to Innovate

	Dependent variable:
	Willingness to Innovate
Trust	0.033 (0.095)
Matching	-1.428 (0.873)
Trust*Matching	0.040 (0.131)
Risk attitude	-0.006 (0.011)
Average trustworthiness	-0.086* (0.049)
Puzzle task ability	0.006 (0.008)
Self-perceived Creativity	0.017 (0.044)
Gender: Male	-0.392 (0.367)
Age	-0.011 (0.014)
Educational attainment: Less than Secondary School Diploma	0.610 (1.074)

Educational attainment: Master of Science	0.780 (0.590)
Educational attainment: PhD	1.215 (1.478)
Educational Attainment: Secondary School Diploma	0.526 (0.404)
Constant	-0.073 (1.974)
Observations	154
Log Likelihood	-93.861
Akaike Inf. Crit.	215.723

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

We then proceeded to introduce two additional controls, that is, *Expectations about Partner's choice* and *Expectations about Partner's success*. In particular, we were interested in the first variable, which is a dummy equal to 1 if the participant believed that the partner had played the DAT, 0 if the participant believed that the partner had played the puzzle task. Its regressor is positively correlated with Willingness to Innovate and significant at 0.01. This means that participants did consider the strategic nature of the interaction and were more likely to play the creative task when they expected the partner to do the same. In the regression we drop Matching and Trust*Matching due to perfect multicollinearity, as Expectations about Partner's choice and Expectations about Partner's success are only defined when Matching is equal to 1.

Interaction and Willingness to innovate

	Dependent variable:
	Willingness to Innovate
Trust	0.127 (0.111)
Risk attitude	-0.010 (0.019)
Average Trustworthiness	-0.100 (0.083)
Self-perceived creativity	0.114 (0.082)
Age	0.001 (0.027)
Puzzle task ability	0.002 (0.012)
Expectations about partner's choice	1.417** (0.700)
Expectations about partner's success	0.510 (0.995)

Educational attainment: Less than Secondary School Diploma	15.515 (1,455.398)
Educational attainment: Master of Science	-0.047 (1.277)
Educational attainment: PhD	1.776 (1.776)
Educational attainment: Secondary School Diploma	1.080* (0.653)
Gender: Male	-0.740 (0.680)
Constant	-4.784 (3.768)
Observations	76
Log Likelihood	-36.621
Akaike Inf. Crit.	101.242

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

Given this result, it is important to understand which variables affected participants' expectations about the partner's choice. Therefore, we decided to run some logistic regressions whose dependent variable was *Expectations about the Partner's Choice.* Firstly, we wanted to understand whether trust played a role in determining these expectations, but we did not find evidence of correlation. We then tried to add other controls, but again no coefficient seemed to be significant. We are thus led to believe that participants' belief that the partner would have followed them in innovation depends on variables that were not included in our dataset. Further research on this topic is needed, because these factors are likely to play an important role in defining not only the expectations about the partner's behavior, but also each participant's choice about innovation.

Trust and Expectations about the partner's choice

	Dependent variable:
	Expectations about partner's choice
Trust	-0.069 (0.081)
Constant	-0.036 (0.511)
Observations	76
Log Likelihood	-50.614
Akaike Inf. Crit.	105.228

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

Regressors and Expectations about the partner's choice

	Dependent variable:
	Expectations about partner's choice
Trust	-0.064 (0.273)
Risk attitude	0.011 (0.053)
Average trustworthiness	0.141 (0.284)
Puzzle task ability	0.002 (0.027)
Expectations about partner's success	-23.965 (6,827.494)
Expected success in the DAT	0.226 (0.768)
DAT score	0.378 (0.355)
Age	-0.011 (0.106)
Gender: Male	-2.107 (2.330)
Educational attainment: Less than Secondary School Diploma	20.284 (10,754.010)
Educational attainment: Master of Science	-20.477 (10,754.010)
Educational attainment: PhD	-5.658 (12,738.270)
Educational attainment: Secondary School Diploma	-0.061 (1.795)
Constant	-7.335 (6,827.541)
Observations	21
Log Likelihood	-7.513
Akaike Inf. Crit.	43.026

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

To conclude, although we confirm the hypothesis that trust would not have a significant effect on Willingness to Innovate when considering both treatments, we do not find enough evidence to infer a correlation between Trust and Willingness to Innovate in the Matching

treatment. The weak effect of Trust did not seem to counterbalance the negative effect of assignment to the Matching treatment on Willingness to Innovate.

2.5.3 Hypothesis 3

To conclude, we wanted to check for a possible nonlinear effect. The literature shows that after a certain threshold the effect of trust on innovation could become weaker or even negative. We ran a logistic quadratic regression, which also includes squared values of the regressors and the variable *Individual*, which is *1-Matching,* that is, 1 if the participant has been assigned to the Individual treatment, 0 otherwise. The variable *Matching* is not squared, because it is a dummy variable whose values can only be 0 or 1. We did not include the variable Trust because we wished to specifically study the nonlinear effect of trust in the two treatments. Moreover, including the Trust variable would have caused issues of collinearity.

The result is the following:

$$Willingness\ to\ Innovate_i = \alpha + \beta Trust_i * Individual_i + \gamma Trust^2{}_i * Individual_i + \delta Matching_i + \zeta Trust_i * Matching_i + \eta Trust_i^2 * Matching_i + \varepsilon_i$$

In this regression, the effect of the interaction term Trust*Matching is therefore:

$$\zeta + 2\eta\ Trust_i * Matching_i$$

ζ is the rate of change of Willingness to Innovate when Trust*Matching is equal to 0, while η tells the direction and steepness of the curve that represents the relationship between the two variables. If η is negative, the curve is concave, if it is positive, the curve is convex. The results of the regression are the following:

Nonlinearity of Trust

	Dependent variable:
	Willingness to Innovate
Trust*Individual	-1.307** (0.612)
Trust^2 *Individual	0.098** (0.045)
Matching	-6.810*** (2.415)
Trust*Matching	0.799 (0.531)
Trust^2 *Matching	-0.059 (0.040)
Constant	3.648** (1.832)
Observations	154
Log Likelihood	-94.573
Akaike Inf. Crit.	201.147

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

The results do not provide any evidence of a significant nonlinear effect of Trust*Matching on Willingness to Innovate, even if the coefficient ζ is positive, while the coefficient η is negative. Instead, in the Individual treatment we can see a significant negative value for the coefficient of the interaction term and a significant positive value for the coefficient of the squared interaction term. Technically, we can infer that in the Individual treatment the relationship between Trust and Willingness to Innovate seems to be nonlinear and best represented by a convex curve with a local minimum. This would mean that for low levels of Trust the effect on Willingness to Innovate is negative, but after a certain threshold of Trust the effect turns weaker or even positive. This would reverse our Hypothesis 3, but more research is needed to investigate the robustness and the causes of this result. To provide additional evidence to the results of the regression, we decided to plot the probabilities that Willingness to Innovate is equal to 1 on both Trust*Matching and Trust*Individual, removing the cases where the latter variables are equal to 0 because the subject does not belong to the corresponding treatment. The plots seems to confirm a U-shaped relationship between Trust and Willingness to Innovate in the Individual treatment. An inverted U-shaped relationship seems to best represent the relationship between Trust and Willingness to Innovate in the Matching treatment. This would confirm Hypothesis 3, but we need to bear in mind that all coefficients related to Trust*Matching in the quadratic regression are not significant. Therefore, we do not find enough evidence to confirm Hypothesis 3.

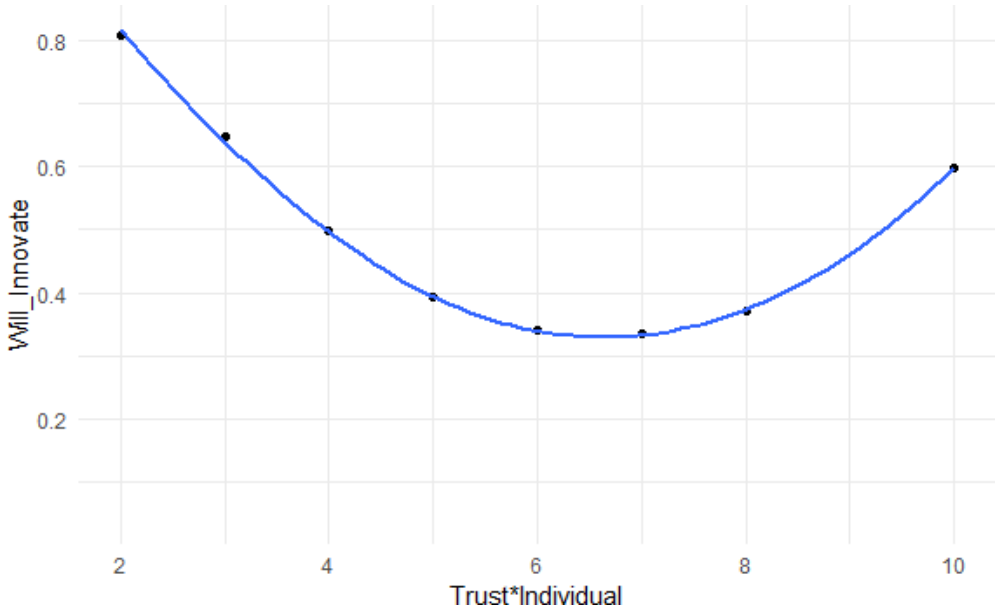


Figure 6: The possible nonlinear effect of Trust on Willingness to Innovate in the Individual treatment

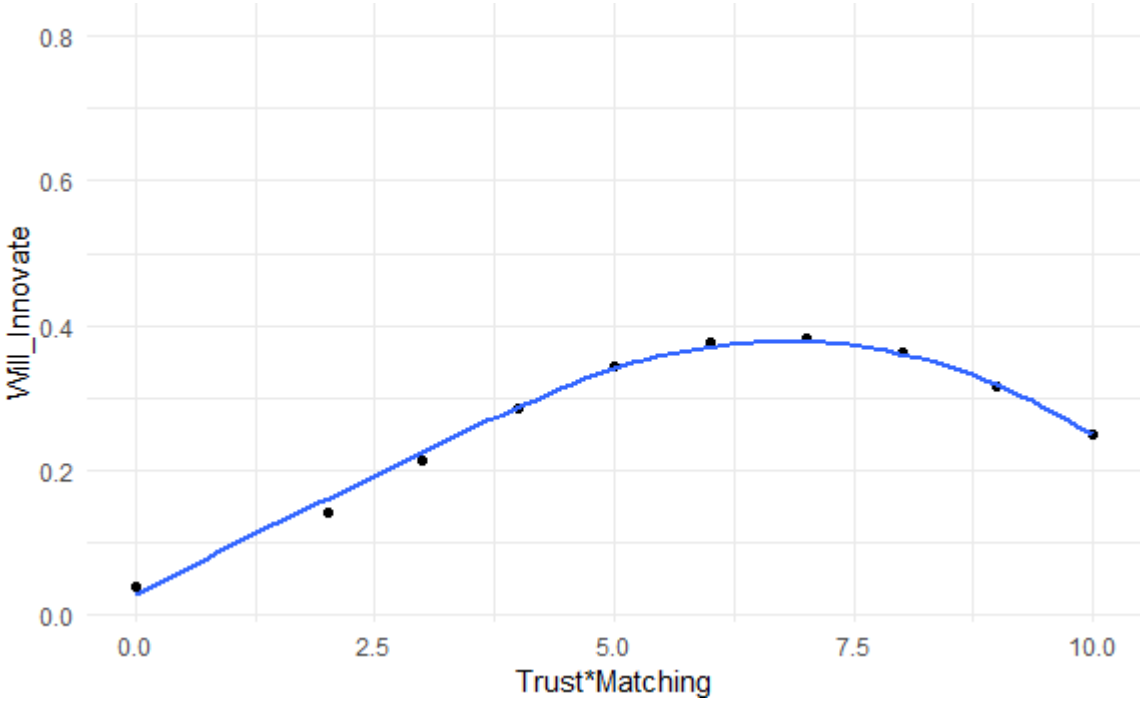


Figure 7: The possible nonlinear effect of Trust on Willingness to Innovate in the Individual treatment

2.6 Discussion

Our work aimed at investigating the existence of a relationship between social trust and willingness to innovate among members of a dyad. To reach this goal, we developed an online experiment on Prolific. Having elicited participants' trust in others and risk attitude, we invited them to play two rounds of a puzzle task and then offered them the opportunity to choose between playing another round of the same task or attempting a new "creative task". In this second round, we created two treatments: the Matching treatment, where pairs were formed, and the Individual treatment. The choice between the puzzle task and the creative task was designed to resemble the choice between exploitation and exploration in an organizational setting (March, 1991), with the former being a safer and more "routinary" option and the latter being riskier but potentially more rewarding. The choice was thus seen as a proxy of participants' willingness to innovate.

We formulated three hypotheses. First, willingness to innovate would be higher in the Individual than in the Matching treatment. Secondly, while the effect of trust on willingness to innovate would not be significant in the whole sample, it would be positive and significant in the Matching treatment. Thirdly, the relationship between trust and willingness to innovate in the Matching treatment would not be linear.

To test our hypotheses, we ran a series of logistic regressions and a Chi-squared test for independence. The results do not provide enough evidence to support all our hypotheses and leave room for further research. The most relevant variables correlated with the choice to attempt innovation seem to be assignment to the Matching treatment and the expectation that the partner had also innovated. The former is negatively correlated with willingness to innovate, while the latter is positively correlated with willingness to innovate. Trust does not seem to have a significant effect neither in the Matching nor in the Individual treatment, and

we do not find evidence of a nonlinear correlation between trust and willingness to innovate in the Matching treatment.

Our results could possibly be explained by variables that we did not include in our study. First, participants in the Matching treatment act in a context characterized by ambiguity (Li et al, 2019; Delle Foglie, Papa and Spagnolo, 2025) and strategic uncertainty (Bruttel et al., 2023). In fact, when participants choose between the puzzle task and the new task, they do not know with which probability the partner will follow them in innovation, thus generating ambiguity. Moreover, strategic uncertainty is determined by the fact that participants' payoff depends on the partner's choice. At the same time, participants do not know the distribution of "types" of partners, that is, how many possible partners in the sample are trustworthy and willing to innovate. Moreover, the nature of the new task is not known when participants decide whether to attempt it. It must be noted that all subjects have already experienced strategic uncertainty and ambiguity and the risk of betrayal in the trust game. Participants in the Matching treatment are thus affected twice by these conditions. Although they do not receive the result of the trust game before the end of the experiment, participants may recall unpleasant feelings related to ambiguity, the risk of betrayal and strategic uncertainty from the trust game. It must be noted that Li, Turmunkh and Wakker (2019) advise against measuring trust only through the trust game, as noise from ambiguity and strategic uncertainty can affect the result, while surveys only capture beliefs of expected trustworthiness. A combined measure of trust could be the best choice, and we decided to follow this advice in the third chapter. As regards the choice in the second stage of our experiment, the more subjects are averse to ambiguity and strategic aversion, the more they could be likely to choose the safer option, that is, not to innovate. This could explain why participants in the Matching treatment were less likely to innovate than participants in the Individual treatment.

Another variable that should have been considered is ambiguity-generated insensitivity, that is, the lack of discriminatory power of subjects in differentiating the degree of likelihoods of ambiguous events (Henkel, 2024). Ambiguity insensitive subjects tend to perceive events that could have different probabilities as equally likely, they are unable to discriminate between composite and single events and do not react to changes in the likelihood of events. In our task, ambiguity-insensitive subjects could believe that the probability that the partner would follow them in innovation was 50%. This would lead them to act independently of their beliefs. For example, high-trust individuals could be less likely to cooperate on innovation than they would have been if they were not ambiguity-insensitive, while low-trust individuals could cooperate more. High levels of ambiguity-insensitivity in the sample could thus reduce the relevance of attitudes like trust or first-order beliefs about the partner in our experiment. Results might depend on how types of subjects are distributed. For example, if most of the ambiguity-averse subjects were concentrated in the Matching treatment, this could explain why the innovative task was rarely chosen. This is particularly true in a context where one's payoff depends on the choice of the partner, and the distribution of types is unknown.

Ambiguity aversion is traditionally measured using the Ellsberg paradox (1961), whereby participants are presented with two urns, each containing 100 balls. One urn contains, for example, 50 black and 50 red balls, while in the second the distribution is unknown. Subjects are asked to bet on a color being extracted from one urn. Ambiguity-averse subjects would choose the urn with a known distribution of colors, to avoid the ambiguity of not knowing the

distribution of colors in the second urn. An alternative approach has been developed by authors such as Li, Turmunkh and Wakker (2019) and Delle Foglie and Papa (2024). The approach is based on the matching probabilities, that is, the probability according to which a subject is indifferent between a prospect with known probabilities, that is, a risky prospect, and a prospect with unknown probabilities, that is, an ambiguous prospect. Specifically, subjects have to face 24 decision situations in which they can choose between a prospect that pays € 15 with probability q and 0 otherwise, that is, a risky prospect, or a prospect where they would get € 15 if the partner decided one of three possible allocations (R, M or S) between her and the subject and 0 otherwise, that is, an ambiguous prospect. If the subject chooses the ambiguous option, the winning probability in the next decision increased, otherwise the winning probability decreased. Ambiguity aversion, ambiguity-generated insensitivity and ambiguity-neutral first-order beliefs, that is, the beliefs of a twin of the decision maker who is perfectly equal to the latter with the exception that she is ambiguity-neutral, are calculated as indexes originated from the average matching probabilities for the three possible allocations chosen by the trustee, either considered as single events or composite events (i.e., the union of M and R). Details can be found in Li, Turmunkh and Wakker (2019). As regards strategic uncertainty, Bruttel et al. (2023) develop an experiment in which participants play a stag-hunt and a market-entry game, where their payoff depends on the unknown strategic choice of the partner. Each game is paired with lotteries that replicate the same payoff structure, but with given exogenous probabilities (i.e., risky lotteries) or with known outcomes but unknown probabilities (i.e., ambiguous lotteries). In each of the three conditions the certainty equivalent, that is, the amount of money players is willing to accept to avoid playing the game or lottery, is calculated. By comparing the certainty equivalents in the three conditions, attitude towards strategic uncertainty is measured. These measures can be used for future research, but possibly in a simplified version to minimize the cognitive load of the experiment.

Participating in a strategic interaction in the Matching treatment leaves participants in a vulnerable position, as the partner could decide to "betray" him by unilaterally deviating from innovation. In an organizational setting, this would mean leaving other team members to bear all costs and risks of innovation with a smaller chance of success, as innovation often requires a high degree of coordination within the team. Delle Foglie, Spagnolo and Papa (2025) prove that, even controlling for ambiguity aversion and ambiguity-insensitivity, betrayal aversion significantly reduces the amount transferred by participants in a trust game. More specifically, participants transfer less when they are matched with a real partner than when they are matched with a machine that randomly decides how much to return to the trustor. In fact, only a human trustee can intentionally decide to return a low amount, and this can be interpreted as a betrayal. We believe that in our experiment betrayal aversion can play a role not only in the trust game, but also in the choice on innovation. Feelings such as anger (Dunn and Schweitzer, 2025) or fear (Neto, Felipe and Bento Caleiro, 2019), together with biases such as loss aversion and status-quo bias (Morrison and Pots, 2008) could also hinder participants' willingness to innovate. The more "extreme" nature of the payoff structure of the DAT, which is more rewarding than the puzzle task in case of success but less rewarding in case of failure, might have triggered extremeness aversion.

We believed that high social trust could have helped mitigate the negative effect of assignment to a treatment that involved a strategic interaction. High-trust people could have been more

willing to accept vulnerability and less scared of being exploited by the counterpart (De Cramer, 1999), thus accepting to take risks and try to coordinate with others. Moreover, high-trust people might have wished to signal their trustworthiness by deciding to innovate, or they could innovate because they thought that other subjects were also trustworthy and innovators, following the self-similar reasoning (Li, Turmunkh and Wakker, 2019).

However, in our work we do not find evidence of a significant effect of trust on willingness to innovate, regardless of which treatment the participant was assigned to. In the Matching treatment, trust has a positive effect on willingness to innovate, but the coefficient is not significant. In fact, the possible positive effect of trust is not enough to counterweigh the negative effect of assignment to the Matching treatment.

However, we do find evidence that, while choosing, players in the Matching treatment understood the strategic nature of the interaction. In fact, participants were significantly more likely to attempt the creative task when they believed that the partner had done the same, because they knew that in the opposite case they would have suffered a loss. We sought to understand what drove participants' beliefs about the partner's choice, but none of the variables included in the experiment appears to be significant, including trust. Further research might help understand which psychological, social, or cultural motives play a role in shaping participants' expectations.

Evidence about a possible nonlinear relationship between trust and innovation is mixed. Opposite to our hypothesis, the nonlinear relationship seems to be present in the Individual treatment instead of the Matching treatment. Moreover, this relationship recalls a U-shaped curve, where under a certain threshold of trust the effect on willingness to innovate was negative, and then it turned positive. Instead, we would have expected an inverted U-shaped curve in the Matching treatment, but we found no significant results to support this hypothesis. The negative effect of assignment to a treatment which involves a strategic interaction seems to be main driver of participants' choice between exploration and exploitation.

To conclude, we want to highlight the limitations of our study. First, variables such as ambiguity, strategic aversion, ambiguity-generated insensitivity and betrayal aversion should be included in the analysis in future studies. The same applies for psychological variables such as fear or anger. Second, of the 208 participants involved in the first stage, only 154 also took part in the second stage. This attrition rate can be justified by the fact the experiment was conducted on two different days, with four days between the first and second stage. A certain degree of attrition was inevitable, but we decided to keep a temporal distance between the two stages of the experiment because we feared that playing the trust game in the same day could have primed participants when choosing between the puzzle task and the creative task. The balance between avoiding the risk of priming and minimizing attrition is difficult to find. Third, the study could be extended to non-British residents, possibly by creating a version of the DAT where participants play in their mother tongue, if scores will still be quickly calculated. This would allow for more variability in our sample, as people from different countries might have different levels of trust, risk attitude and willingness to innovate, and they could thus show a different behavior in a strategic interaction setting. In this case, one should also introduce controls for social and cultural factors that may explain this different behavior. To conclude, we could change our experiment by informing participants of the

partner's level of trust and create non-random pairs based on reciprocal trust levels. Knowing the partner's characteristics could induce feelings of similarity or dissimilarity with the partner. It would thus be possible to investigate for motives explaining participants' expectations on the partner's behavior, and to understand the role played by expected trustworthiness. Creating non-random pairs could help identify possible differences in behavior between homogeneous and heterogeneous dyads, thus extending the scope of our research.

2.7 Conclusions

Our experiment aimed at investigating whether trust affected willingness to innovate among members of a dyad involved in strategic interaction. In the two-stage experiment that we designed, however, we do not find evidence of a significant effect of social trust in participants' choice between exploration and exploitation, that is, between the new creative task and the well-known puzzle task. This holds true both for the Matching and the Individual treatment. Instead, it seemed that being assigned to the Matching treatment, which involves a strategic interaction, hinders willingness to innovate, while expecting that the partner had innovated favors innovation. Trust could not counterweigh the negative effect of assignment to the Matching treatment. Omitted variables such as ambiguity aversion, ambiguity-generated insensitivity, betrayal aversion and fear could have driven participants away from innovation when matched with a partner. The fact that participants were more likely to innovate when they expected the partner to do so reveals that they understood the strategic nature of the choice between the puzzle task and the DAT, but we could not identify the factors driving participants' expectations on the partner's choice. No evidence of a nonlinear relationship between trust and willingness to innovate is found in the Matching treatment, while in the Individual treatment evidence seems to suggest a U-shaped relationship. Further research might help shed light on the motives driving the latter phenomenon, which seems to be in contrast with our hypotheses and the literature. Our study- a two-stage experiment run on Prolific- has some limitations which could have had an impact on our results. Attrition rate and the excessive homogeneity of the sample could at least partly explain the lack of significance in the experiment, while providing information on the partner's level of trust could help reduce the level of ambiguity and strategic uncertainty.

3 Trust, Partner's Trust and Willingness to Innovate in a strategic setting

3.1 Introduction

The aim of this work is to further investigate the relationship between trust and innovation in dyads engaging in a strategic interaction. Moreover, we wish to understand whether receiving information on the partner's level of trust influences subjects' willingness to innovate. To reach our goal, we built a two-stage experiment with students from the University of Trento, whereby the choice about innovation was modelled as the choice between a well-known task and a new task, which was riskier, but potentially more rewarding.

As we underlined in the first two chapters, the importance of trust and innovation for economic and social development has been stressed by many authors in the last century (e.g., Knack and Keefer, 1997; Zak and Knack, 2001; Solow, 1956, Romer, 1986). While being both multifaceted concepts with many possible definitions, trust and innovation shared the common features of risk, ambiguity, uncertainty and a positive expectation about either another person or future gains. Trust involves ambiguity as the probability of the partner's trustworthiness are unknown. Moreover, trust implies vulnerability to deception and betrayal by the counterpart, especially when the latter is a stranger. Innovation is affected by uncertainty about the result and quality of the innovation, the ability to outperform competitors whose strategic moves cannot be fully understood and to absorb innovation from the market. As we will briefly summarise in the literature review section, there is ample evidence of a positive correlation between trust and innovation, even if excessive level of trust could have a weaker or even negative effect on innovation.

We built a controlled online experiment with students from the University of Trento to further investigate the relationship between trust and willingness to innovate. As outlined in the second chapter, using experimental techniques constitutes a novelty with respect to the literature. In fact, most existing works are either observational studies or surveys. Through experiments, we can create a controlled environment where it is easier to distinguish between exogenous and endogenous variables and to make causal inferences. Experimental measures of variables such as trust or risk attitude can be elicited simulating real-life conditions and avoiding social desirability bias. Random assignment of subjects to treatment conditions allows us to overcome selection bias, and information provided to subjects can be easily controlled. Moreover, experiments are easier to replicate for future research. Another novelty of our experiment is the choice to nest the decision to innovate within a prisoner's dilemma condition based on March (1991)'s exploration-exploitation dilemma. Participants got used to a task by playing it five times, then they were matched with a partner and asked whether they wanted to play another round of the same task or a new task, which was riskier, but potentially more rewarding. We can describe this choice as one between exploitation, that is, refining existing technologies and processes, and exploration, that is, searching for new solutions. Exploitation guarantees safer but lower gains, while exploration implies more "extreme" gains, which are lower than exploitation in case of failure, but higher in case of success. Organisations usually prefer exploitation over exploration, unless they have strong positive expectations that exploration will be successful. In our experiment, dyads gain the highest joint payoff if they coordinate on choosing the new task. However, as the payoff structure is modelled after a prisoner's dilemma, defecting and continuing with the known

task always maximizes individual payoff. The only Nash equilibrium of the game is reached when both players defect. However, as we will show in the literature review section, there is ample evidence that subjects do cooperate in a prisoner's dilemma setting, even when the game is one-shot. The most frequent explanations for cooperative behavior in social dilemmas are unconditional altruism, be it pure (Anderson, 1998) or motivated by a positive utility from giving to others (Andreoni, 1988), and reciprocity, that is, the willingness to help those that are kind and punish those that are unkind (e.g., Rabin, 1993; Hayashi, 1999; Kiyonari, Tanida and Yamagishi, 2000; Cox, 2004). Other motives include personality traits such as prosociality (Van Lange and Kuhlman, 1994), high self-monitoring or sensation seeking (Boone, De Brabander and van Witteloostuijn, 1999), high conscientiousness, extroversion and agreeableness (Le Pine and Van Dyne, 2001) or low neuroticism (Ben-Ner, Putterman and Kong, 2004; Schroeder, Nettle and McElraeth, 2015). Feelings such as guilt aversion (Battigalli and Dufwenberg, 2007) and the desire to send a signal of trustworthiness, sometimes motivated by the feeling that other subjects are also trustworthy (Li, Turmunkh and Wakker, 2019) could also play a role. Moreover, social preferences such as inequality aversion (Schneider and Schields, 2022; Cox, 2004; Di Bartolomeo and Papa, 2016a) or even a simple fairness heuristic (Capraro, Jordan and Rand, 2014) can explain cooperation. We believe that trust could be another personality trait that favours cooperation, as studies in the literature show that its effect can be distinguished from that of unconditional altruism (Cox, 2004; Di Bartolomeo and Papa, 2016a). In fact, trust could help participants overcome the fear of being exploited by the counterpart (De Cremer, 1999) and decide to try a new task, knowing that the partner will do the same. Moreover, Chaudhuri, Sopher and Strand (2002) find that cooperation in the prisoner's dilemma is correlated with higher trust in the trust game. Therefore, we hypothesize that Trust will have a positive effect on Willingness to Innovate.

With respect to the second chapter, we introduced new features in our experiment. First, when dyads were formed, participants received information on the partner's level of trust- whether it was high or low- and whether it was similar to their own. Different treatments were created, whereby high-trust, low-trust or mixed pairs were created. In an experimental context characterized by ambiguity (Li, Turmunkh and Wakker, 2019) and strategic uncertainty (Bruttel et al., 2023), knowing the partner's level of trust could help changing subjects' subjective perception regarding the probability distribution of being matched with a collaborative mixed or selfish type, thus facilitating the choice to innovate. In fact, as we will explain in detail in the next section, the literature shows that providing information on some personality traits of the partner can increase subjects' willingness to cooperate. For example, subjects who know or perceive that their partner scored high in agreeableness and openness to experience (Guilfoos and Kurtz, 2017), and honesty (Van Lange and Kuhlman, 1994), that she has a higher probability of reciprocating her choice (Acevedo and Krueger, 2005), she is more similar to them (Fisher and Savranevski, 2023; Li, Turmunkh and Wakker, 2019) or belongs to the same group (e.g., De Cremer and Van Vugt, 1999; Ahmed, 2007; Yamagishi and Kiyonari, 2000), could cooperate more in a social dilemma. We hypothesize that receiving the information that the partner is a high-trust type will foster Willingness to Innovate. Based on the literature, we provide two possible explanations for our hypothesis. First, knowing that the partner is a trusting person could induce subjects to believe that that they will be reciprocated if they decide to cooperate on innovation. The strategic uncertainty and ambiguity about the probability of a cooperative choice by the partner would thus be reduced. As found by Acevedo and Krueger (2005), manipulating *projection,* that is, subjects' perceived

probability of reciprocation, could increase cooperation. Alternatively, information on the partner's Trust could be seen as a proxy of her level of morality or his score in other personality traits that lead participants to believe that their partner is a good person. Participants could thus feel urged to adhere to social norms such as inequity aversion or perceive guilt aversion (Battigalli and Dufwenberg, 2007) and thus desire not to disappoint the partner. Both channels would lead to higher cooperation and thus higher Willingness to Innovate in our experiment. Moreover, we hypothesize that the impact of information on partner's trust could be stronger for participants with low trust. In fact, the latter could change their mind about cooperation when knowing that their partner scored high on Trust, while high-trust participants could be already more inclined to cooperate.

Other novelties that we introduced in this third chapter include the measurement of trust, the length of the training phase before the choice and the nature of the new task. We decided to choose a combined measure of trust, obtained as the average between the score in the trust game and the answer to the World Values Survey's Rosenberg question. In fact, while survey measures of trust are potentially subject to social desirability bias, they are not subject to the noise generated by ambiguity and strategic uncertainty which characterise the trust game (Li, Turmunkh and Wakker, 2019). In fact, the trust game is a strategic task played with an unknown partner who could decide to betray with an unpredictable probability. As regards the training phase, we decided to extend it by asking participants to play the puzzle task five times. This reinforces the lock-in effect (Foxon, 2014) by creating a stronger habit. To conclude, we changed the new task that players could choose in alternative to the puzzle task. We opted for a task which is more similar to the puzzle task, but it has a completely different solution strategy. This choice was made to avoid possible confusion originating from the excessive difference between the Divergent Association Task (DAT) and the puzzle task in the experiment that we ran in the second chapter. However, the different solution strategy guarantees that the new task still represents an explorative choice with respect to the puzzle task.

The experiment, including the hypotheses, the design plan, the sampling plan, the variables and the analysis plan, was preregistered on the OSF platform before the first experimental session.

Result partially confirm our hypotheses. No significant correlation is found between trust and willingness to innovate, and this remains true even excluding participants assigned to the Baseline and thus not paired with a partner. Instead, the effect of information about the partner's trust on willingness to innovate is positive and significant. Knowing that the partner is a high-trust type increased the likelihood that participants chose the new task. Evidence showed that a possible channel of this positive correlation could be participants' belief about the partner's choice. In fact, participants who got to know that their partner had scored high in trust seemed more likely to believe that the partner had chosen the new task. However, evidence in favor of this channel is not particularly strong, suggesting that other mediating variables could have come into play. To conclude, we find no evidence of a different effect of partner's trust between high-trust and low-trust individuals.

Our paper proceeds as follows. In the first section we reviewed the literature on Trust and Innovation and explained our contribution. In the second section we listed and justified our

research hypotheses. In the third section we described and motivated the experimental design in detail. In the fourth section we presented our results, which we discussed in detail in the fifth section. The last section was dedicated to conclusive remarks.

3.2 Literature review and contribution

In the first two chapters, we highlighted the importance of trust (e.g., Knack and Keefer, 1997; Zak and Knack, 2001; Rothstein and Uslaner, 2005; Hunt, 2012) and innovation (e.g., Solow, 1956; Romer, 1986) in fostering economic and social development in modern society. We described how behavioral economics provides tools to investigate why people trust others even in social dilemmas (e.g., Evans and Krueger, 2009; Dunning, Fecthenhauer and Schlösser, 2012; Zak, 2006; Di Bartolomeo and Papa, 2016b) and it helps explain why organizations often struggle with innovation (e.g. Pots, 2010; Neto, Felipe and Bento Caleiro, 2019). The literature that we examined in the first chapter mostly identifies a positive correlation between trust and innovation. High-trust societies tend to have better innovation outcomes (e.g., Dakhli and De Clercq, 2004; Thompson, 2018; Lyu, Ma and Zhang, 2022), trust fosters innovation in teams (e.g., Clegg et al., 2022; Shazi, Gillespie and Steen, 2014), innovation involving cooperation between different organizations is facilitated by trust (e.g., Gulati, 1998; Brockman et al., 2018; Steinbruch, 2021) and firms acting in a high-trust external and internal environment tend to outperform rivals in innovation (e.g., Watson Wyatt, 2002; Meng et al., 2019; El Ghoul et al., 2023). Among the most cited explanations, trust is shown to facilitate knowledge sharing, reduce monitoring and transaction costs, mitigate the fear of exploitation by the partner and foster the proposal and implementation of new ideas. However, it has also been shown that trust has a "dark side", leading to a nonlinear effect on innovation. In fact, excessive levels of trust can dangerously reduce monitoring and constructive task conflict (Langfred, 2004; Bidault and Castello, 2010) and over-relying on high-trust partners can cause organisations to lose market opportunities (Budunchi, 2013).

Our work involves a decision between innovation and an established habit in a prisoner's dilemma setting. The alternative to innovation is playing another round of a well-known task. The situation that each dyad faces resembles the exploration-exploitation dilemma (March,1991). In fact, they choose between a safer and more familiar task and a riskier, but potentially more rewarding alternative task. Moreover, we nest the exploration-exploitation dilemma in a prisoner's dilemma setting, where cooperating means attempting to innovate and defecting means sticking to a habit. This is meant to reproduce a situation where team members can decide together whether to innovate. They have divergent opinions and characters, and they know that they can grasp the full benefits of innovation only if they innovate together. Innovation is thus envisioned as a coordination issue, whereby participants can decide whether to coordinate with the partner to innovate or defect to follow a well-established habit. The Nash equilibrium is obtained when both players decide to defect. Defection is a dominant strategy, that is, it is the most profitable strategy for every possible action of the partner. Ahn et al. (2001) identify greed and fear of opportunistic behavior by the counterpart as the main factors that drive participants away from cooperation in strategic dilemmas. Li, Turmunkh and Wakker (2019) and Delle Foglie, Papa and Spagnolo (2025) underline the role of ambiguity aversion, strategic uncertainty and betrayal aversion in hindering cooperation in strategic settings. In fact, in our setting participants do not know the task that they will play and the probability with which the partner will cooperate for innovation. Moreover, by cooperating, participants become vulnerable to betrayal by the

counterpart, and this thought often hinders cooperation. In our experiment, information about the partner's trust partially reduces the uncertainty regarding the type the partner belongs to in terms of trust.

However, there is ample evidence of cooperation even in one-shot strategic dilemmas. In this section we will expand on the results described in the first chapter and in the literature review of the second chapter, analyzing other studies on the drivers of cooperation. In particular, we will focus on expectations on the partner and individual traits.

Van Lange and Kuhlman (1994) maintain that individuals can apply two different kinds of rationality to social dilemmas. If they apply individual rationality, they believe that it is rational to defect, because it always yields the best outcome for oneself. If they apply collective rationality, they believe that it is better to cooperate, as cooperation yields the highest joint payoff. In terms of Social Value Orientation (SVO), individuals can be classified as prosocials, competitors and individualists. The authors measure participants' SVO characteristics, then they inform them of the partner's morality and intelligence scores and ask them to play different versions of a social dilemma. They find that prosocials expect more cooperation from the partner and cooperate more than individualists and competitors. More cooperation is expected from high-honesty partners, particularly by prosocials, and from low-intelligence partners, particularly by individualists and competitors. As regards actual cooperation, participants paired with high-honesty partners cooperate more, especially if they are prosocials. At the same time, all participants paired with high-intelligence partners cooperate more. Paradoxically, competitors and individualists cooperate more with high-intelligence partners, although they expect them not to reciprocate. One possible explanation is perceived similarity to the high-intelligence partner. In fact, it has been shown that players are more likely to cooperate if they perceive the partner as similar to them. (Fischer and Savranevski, 2023). The symmetry argument is known as "the basic argument in favour of cooperation", and it asserts that rational players should understand that they are in the same situation as their partner and choose to cooperate. However, as the distribution of prosocial and proself types in the population is unknown, the perception that others are similar to self is often wrong. This behavioral bias is known as the false consensus effect (Ross, Greene and House, 1977) and it often leads cooperators to lose money in strategic games (Li, Turmunkh and Wakker, 2019). Fischer (2009) introduces different methodologies to induce expectations of similarity. The result is that players cooperate more in tasks with a lower similarity threshold, that is, the similarity level that must be reached to make cooperation more convenient. Moreover, people with higher perceived similarity induce more cooperation. Comparable results can be found for social identification, that is, the level of kinship with a group. De Cremer and Van Vugt (1999) find that social identification promotes cooperation in a social dilemma. The effect is particularly strong for proself subjects, who redefine their goal from personal to group welfare. Similarly, Ahmed (2007) finds that subjects matched with partners from the same group cooperate more in the prisoner's dilemma and the public goods game. The effect comes more from in-group favoritism rather than hostility towards partners from different groups. Yamagishi and Kiyonari (2000) found that the driver of favoritism towards in-group partners is expectations of higher generalised reciprocity with respect to out-group partners. The effect is seen only if all partners know that they belong to the same group. Instead, Ciccarone, Di Bartolomeo and Papa (2020) find that in-group favoritism is not

caused, neither directly nor indirectly, by beliefs on the partner's behavior, but it is solely caused by an intrinsic preference for group members.

Another approach comes from Acevedo and Krueger's (2005). The authors propose two theories on how participant form beliefs on their partner's future action. People can either anticipate the likelihood with which their action could be reciprocated, a process called *projection*, or formulate hypotheses on the partner's most likely action and act consequently, a process called *assimilation.* Acevedo and Krueger attempt to manipulate expectations resulting from both projection and assimilation. The result is that only manipulating projection is correlated with higher rates of cooperation. Guilfoos and Kurtz (2017) find that participants who receive information that their partner scores high in agreeableness and openness to experience are more likely to play cooperatively and expect more cooperation from the partner in the public goods game. Agreeableness and openness to experience are two of the *Big Five Personality Traits,* the other three being neuroticism, extroversion and conscientiousness. Al-Ubaydli, Jones and Weel (2015) find that a pair's average cognitive ability enhances joint cooperation in a repeated prisoner's dilemma. Only openness to experience predicts cooperative behavior in the first round. Perceiving the other member of team as a stranger or a partner is also likely to affect cooperation. Results are mixed, with Andreoni (1988) and Palfrey and Prisbey (1996) finding that strangers are more cooperative and Croson (1996) and Suneja and Das (2024) finding the opposite result (see Andreoni and Croson, 2001).

Another field of literature studies the effect of different personality traits and social preferences on individual cooperation in a social dilemma. Boone, De Brabander and van Witteloostuijn (1999) find that individuals high in self-monitoring, that is, the ability to regulate their behavior and self-presentation, have higher rates of cooperation in the prisoner's dilemma. Moreover, high sensation seekers, that is, people who perceive a strong need to experience new and varied stimuli, are more likely to cooperate. Studies have been undertaken about people with *internal* or *external locus of control*, that is, people who respectively tend to believe that events in their lives depend mostly on themselves and external factors. Results are mixed. Bialer (1961) finds that external children are more cooperative in a cooperation board game, while Boone, De Brabander and van Witteloostuijn (1999) find that internals are more likely to cooperate in a repeated prisoner's dilemma, because they use cooperation as a signal to induce trust from the partner in the following rounds, and they are more likely to switch to competitive behavior when the risk of retaliation is lower. Many authors have investigated the correlation between the Big Five Personality Traits and cooperation in a social dilemma. For example, LePine and Van Dyne (2001) find that participants who score high in conscientiousness, extroversion and agreeableness and who score low in neuroticism are more cooperative. Ben-Ner, Putterman and Kong (2004) find that among women cooperation is enhanced by agreeableness and hindered by conscientiousness and neuroticism, while among men cooperation is enhanced by agreeableness and conscientiousness and hindered by extraversion. Schroeder, Nettle and Mc Elreath (2015) find that extraversion is associated to more cooperation in the public goods game when the institution of punishment is present, while neuroticism is associated with lower cooperation and agreeableness is associated with higher contributions, but only in the first rounds of each game. To conclude, participants may be afraid of disappointing the partner, or even fear that the partner thinks that they purposedly want to damage her. This

feeling, which is called guilt aversion (Battigalli and Dufwenberg, 2007) originates especially if subjects perceive the partner to be a moral person and can lead to cooperative behavior.

As regards social preferences, two types of models have been presented in the literature (Suneja and Das, 2024). According to altruistic preference models, individuals are willing to cooperate irrespective of the partner's behavior. According to reciprocal preference models, individuals are willing to reward partners who treat them kindly and punish partners who treat them unfairly.

Becker (1974) and Andreoni (1988) find that players enjoy giving contributions to the public good, experiencing what is called a "warm glow" and getting private utility from it. Andreoni and Miller (1993) find that there is a stable fraction of altruists in the population, which continue to cooperate even in repeated single shots of the prisoner's dilemma. Altruism can thus also be "pure" and not influenced by warm glow. The finding that pure altruism exists in a public goods game is confirmed by Anderson (1998). Altruism can be separated from trust (Cox, 2004; Di Bartolomeo and Papa, 2016a) because it is a social preference unconditional on both beliefs about, and the actual choices of, the partner.

Rabin (1993) finds that in social dilemmas people are willing to help those who are kind and punish those who are unkind. Hayashi et al. (1999) confirm the finding by reporting that players in their experiment tended to treat the one-shot prisoner's dilemma as if it was an assurance game and cooperate only when they see or expect cooperation from the partner. Expectations for the partner's behavior are mainly influenced by general trust among American participants and by sense of control among Japanese participants. Kiyonari, Tanida and Yamagishi (2000) confirm Hayashi et al. (1999)'s finding by claiming that participants are influenced by the *social exchange heuristic,* thus they see the prisoner's dilemma as a social exchange, where it is optimal to obtain mutual cooperation. Cosmides and Tooby (1989) and Gintis et al. (2002) maintain that humans have acquired a predisposition to reciprocity throughout evolution. Cox (2004) finds evidence that positive reciprocity plays a role in fostering cooperation in social dilemmas, and it can be distinguished from inequality aversion, as the latter is unconditional on the partner's choice.

Authors have investigated the effects of other social preferences on cooperation in social dilemmas. Capraro, Jordan and Rand (2014) find that a fairness heuristic or the belief that the partner would anchor at a 50% contribution provide a good explanation for players' choices. Instead, Schneider and Schields (2022) suggest that inequity aversion is the most important driver of cooperation in simultaneous prisoner's dilemma. Boone, Declerck and Kiyonari (2010) find that trust is more important to enhance cooperation than explicit incentives among prosocial participants, while the opposite holds for proself participants. Chaudhuri, Sopher and Strand (2002) find that subjects who chose to cooperate in a repeated prisoner's dilemma exhibit higher levels of trust in the trust game, while no effect is found for reciprocity.

In this section, we provided an overlook of the literature on trust and innovation and the literature on personality traits and cooperation in the prisoner's dilemma. Our goal is to bridge these two fields of literature to find how trust and knowledge on the partner's trust affects willingness to innovate. Therefore, we developed a laboratory experiment which simulates a choice between innovation and a well-established habit in a prisoner's dilemma context. Innovation is seen as a coordination issue, whereby the choice between innovating

and sticking to a habit is done simultaneously by members of a dyad. In fact, many innovations happen within a team, whose members must collectively decide whether to follow a new path. Conflict is inevitable in such a scenario, as some team members could be reluctant to challenge well-established habits and risk an innovation. Innovators must be able to convince partners that the risk is worth taking and innovation will bring benefits for all, albeit possibly delayed and less certain than benefits from exploitation. Moreover, innovators need to feel at ease to communicate their ideas to partners without fearing to be ridiculed or even sidelined. This conflict is reproduced in a simplified version in our experiment. The literature seems to provide evidence that many participants will choose to attempt an innovation when they are matched with a partner, even if defecting to the well-known task is a dominant strategy. We aim to study whether high-trust people are more likely to choose to innovate in this strategic setting. Furthermore, we investigate whether information about the partner's trust level—whether higher, lower, or equal to that of the participant—affects willingness to innovate. In the next section we will describe the hypotheses of our experiment.

3.3 Hypotheses

In the previous section we analysed the literature and outlined our contribution. We now proceed to list our hypotheses. As we will explain later, our experiment is divided into two stages. In the first stage, trust and risk attitude were elicited. In the second stage, participants played five rounds of a puzzle building task (Puzzle 1), then they were asked to decide whether to attempt a new task, called Puzzle 2, or to play another round of Puzzle 1. In the second stage, participants were assigned to either the Baseline or one of the treatments. As we will explain later, in the treatment *High-High* high-trust participants were matched with a high-trust partner, in the treatment *Low-Low* low-trust participants were matched with a low-trust partner, in the treatment *Mixed* participants were matched with a partner with a different level of trust. In all treatment conditions, participants were informed of the partner's level of trust and their payoff in the chosen task partially depended on the partner's choice. In the Baseline, instead, no pairs were formed. We expected trust and partner's trust to have a positive effect on willingness to innovate in the treatments. Moreover, we expected the effect of partner's trust on willingness to innovate to be stronger for individuals with low individual trust.

Hypothesis 1: *The effect of Trust on Willingness to Innovate will be non-significant in absence of strategic interaction, positive and significant when strategic interaction happens.*

In the literature review, we have identified risk and positive expectations about the counterpart's actions as the two central pillars of trust. In fact, every interaction leaves individuals potentially vulnerable to opportunistic behavior and betrayal by their counterparts. Nevertheless, people choose to engage in such interactions when they hold at least some confidence that their partner will not deceive them. In our setting, no strategic interaction is present in the Baseline, therefore we do not expect to see any significant effect on Willingness to Innovate. Instead, in the three treatment conditions participants are engaged in a prisoner's dilemma, whereby participants decide simultaneously whether to cooperate for innovation or to defect and continue with the well-known task. Although defection is a dominant strategy, the literature suggests that we will see some degree of cooperation. Trust could be one of the factors contributing to cooperation. In fact, high-trust participants could tend to focus more on the opportunities rather than the risks of attempting an innovation. Behavioral biases such as fear of being ridiculed (Neto, Felipe and Bento

Caleiro et al., 2019), confirmation bias (Neto et al., 2022), status quo bias (Heidenreich and Kraemer, 2015) or loss aversion usually prevent individuals from attempting to innovate. Ambiguity aversion (Li eta al., 2019; Delle Foglie, Spagnolo and Papa, 2025) and aversion to strategic uncertainty (Bruttel et al., 2023) can hinder cooperation, while ambiguity-generated insensitivity can reduce participants' eagerness to act based on their beliefs, as they view all outcomes as equally likely. Therefore, we expect that willingness to innovate will be higher in the Baseline than in the treatments.

However, trust can help individuals overcome behavioral biases (e.g., Covey, 2006; Lyu, Ma and Zhang, 2022) and the fear of being exploited (De Cremer, 1999), thus making innovation appear as a less risky option. Instead, high-trust subjects will believe that their innovative ideas will be listened to, and they will gain a benefit from innovation (Clegg et al., 2022). Secondly, trust can create a social norm (Burnham, 2000; Hermann and Ariely, 2006) which interprets innovation as cooperation and penalizes defection through the risk of reputational losses (Lyu, Ma and Zhang, 2022; Akçomak and ter Weel ,2009). Norms such as inequity aversion or heuristics as the *fairness* or *social exchange heuristic* (e.g., Schneider and Schields, 2022; Capraro, Jordan and Rand, 2014; Kiyonari, Tanida and Yamagishi, 2000) could be more relevant for high-trust participants, thus favoring cooperation. High-trust partners could be more likely to apply collective rationality to the prisoner's dilemma, thus focusing on reaching the highest joint payoff for the dyad and seeing the task in terms of morality rather than might (Van Lange and Kuhlmann, 1994). Moreover, trust could drive expectations of reciprocity by the partner, as reported by Hayashi et al. (1999) for American participants in a prisoner's dilemma. To conclude, we expect a positive effect of trust on willingness to innovate in the three treatments.

To verify Hypothesis 1, we created the dummy variable *Willingness to Innovate,* which is equal to 1 if the participant has chosen Puzzle 2, and 0 if she has chosen Puzzle 1. Moreover, we created the dummy variable *Treatment*, which is equal to 0 if the participant has been assigned to the Baseline, and 1 if she has been assigned to one of the three treatments. An interaction variable between Treatment and Trust was also created. As Willingness to Innovate, a dummy variable, is the dependent variable, we ran a series of logistic regressions to test the hypotheses. These included the following:

$$Willingness\ to\ Innovate_i = \alpha + \beta Trust_i + \varepsilon_i$$

$$Willingness\ to\ Innovate_i = \alpha + \beta Trust_i + \gamma Trust_i * Treatment_i + \delta Treatment_i + \varepsilon_i$$

Hypothesis 2: *Being matched with a high-trust partner has a positive and significant effect on Willingness to Innovate.*

This hypothesis refers to the behavior of dyads in the three treatments. According to the literature, providing information on one of the partner's personality traits could influence participants' choice between cooperation and defection in a Prisoner's Dilemma. Knowing that the partner is a high-trust type could be interpreted as a signal that she is also honest and intelligent, or that she scores high in personality traits such as agreeableness and openness to experience, which could lead to higher cooperation rates (Van Lange and Kuhlman, 1994; Guilfoos and Kurtz, 2017). If the partner is perceived as honest or intelligent, participants could identify with her or experience guilt aversion (Battigalli and Dufwenberg, 2007) and wish not to disappoint her. These would be strong motives to avoid opportunistic behavior and choose to cooperate on innovation. Moreover, the information might affect participants'

expectations of reciprocity, as if the *projection* mechanism was manipulated (Acevedo and Krueger, 2005). The literature suggests that expectations of reciprocity could be one of the main drivers of cooperative behavior (e.g., Rabin, 1993; Hayashi et al., 1999). Participants in social dilemmas tend to cooperate more when they expect that they will be reciprocated. Especially in the *High-High* treatment, knowing that the partner is a high-trust type can generate feeling of expected similarity (e.g., Fischer, 2009) or social identification (e.g. De Cremer and Van Vugt, 2009; Yamagishi and Kiyonari, 2000). Members of the high-trust dyads could believe that they are matched with a partner belonging to the same group, thus triggering in-group favoritism and expectations of reciprocity. If high-trust individuals are more likely to be prosocial and thus tend to see social dilemmas in terms of morality rather than strength (e.g. Van Lange and Kuhlman, 1994), knowing that the partner is high-trust could strengthen social norms and heuristics favoring cooperation.

In the *Low-Low* treatment, instead, knowledge that the partner is a low-trust type could make defection more morally acceptable. In fact, participants will receive a clue that the partner is not likely to reciprocate and cooperate on innovation, and this effect could be particularly strong in individuals whose preference for cooperation could be already lower than average, if Hypothesis 1 is confirmed. Hypotheses related to the *Mixed* treatment will be outlined in Hypothesis 3.

To test these hypotheses, we excluded subjects assigned to the Baseline and we ran a series of logistic regressions. The dependent variable was Willingness to Innovate, while the main regressors were Trust and *Partner's Trust,* a dummy variable equal to 1 if the participants' partner had a trust score higher than the sample median, 0 otherwise. The simplest regression, before controls were added, was the following:

$$Willingness\ to\ Innovate_i = \alpha + \beta Trust_i + \gamma Partner's\ Trust_i + \varepsilon_i$$

Hypothesis 3: *The effect of Partner's Trust on willingness to innovate is weaker for subjects with high individual trust relative to those who have low individual trust.*

Hypothesis 3 is the most challenging to prove, as the literature does not provide solid evidence on what happens in dyads where partners have different levels of trust. We hypothesize that there could be a moderating effect, that is, Trust could either weaken or strengthen the effect of Partner's Trust on Willingness to Innovate. To prove it, it is essential to compare the behavior of dyads in the *Mixed* treatment to the behavior of dyads in the *High-High* and *Low-Low* treatments. One possible explanation might come from De Cremer and Van Vugt (1999). The authors maintain that group identification might increase cooperation due to *transformation of motivation*, whereby self-interested individuals redefine their motivation and aim at maximizing the collective outcome of the group. A similar transformation might happen for subjects with low individual trust when they are informed that their partner is a high-trust type. Low-trust subjects could experience an "identity crisis" and redefine their attitude towards cooperation, knowing that their partner is likely to reciprocate them if they choose to innovate. Moreover, they could develop a positive attitude towards the partner, seeing him as moral and intelligent, and they might even experience a sense of guilt for not being as moral as their partner. These thoughts could lead the low-trust participant to change her strategy and decide to cooperate on innovation. Instead, participants with high trust could be already more likely to cooperate on innovation, according to Hypothesis 1 and Hypothesis 2. When a high-trust participant is matched with a high-trust partner, information

on Partner's Trust does not cause an abrupt change of strategy. When a high-trust participant is matched with a low-trust partner, we believe that a change of strategy is less likely than in the case when a low-trust participant is matched with a high-trust partner.

To test Hypothesis 3, we created the interaction variable between Trust and Partner's Trust, and we ran the following regression:

$$Willingness\ to\ Innovate_i = \alpha + \beta Trust_i + \gamma Partner's\ Trust_i + \delta Trust_i * Partner's\ Trust_i + \varepsilon_i$$

As we will explain in the Results section, our hypothesis is confirmed if the coefficient δ is negative, thus indicating that higher trust weakens the effect of Partner's Trust on Willingness to Innovate.

3.4 Experimental design

To conduct our study, we designed a laboratory experiment with a sample of students from the University of Trento. The optimal sample size was calculated using a z-test for proportions in GPower, setting the power to 0.80 and the confidence level to 0.95. This yielded a required sample of 160, which we increased to 170. According to GPower calculations, the given sample size would have been enough to identify a difference in proportions of around 0.26 between each pair of treatments. The experiment was carried out in nine supervised online sessions. Participants in each session first joined a Zoom meeting and then accessed the experiment simultaneously through a link and an entry label provided in the chat. Each session lasted approximately 30 minutes, and participants could only take part once. Earnings were expressed in Experimental Currency Units (ECU), which were converted into Euros at the following rate:

1 ECU= € 0.01

Each participant gained a show-up fee of 300 ECU, that is, € 3. Moreover, participants could get a bonus payment based on their performance in the experiment. The maximum possible bonus payment was 5.39, leading to a maximum total payment of € 8.39. Only participants who completed the entire experiment were included in the data analysis, except for two subjects who were matched with a partner who did not complete the experiment and were therefore excluded. In total, 6 participants were excluded from our data analysis, reducing the real sample size of the experiment from 170 to 164. The language used in the experiment was Italian.

Our experiment was divided into two stages. Unlike the second chapter, the two stages happened consecutively in the same session, with no time interval. This allowed us to almost nullify the attrition rate. In the first stage, we elicited trust and risk attitude. In the second stage, subjects completed five rounds of a puzzle task, called Puzzle 1. Later, they could choose whether to play another round of Puzzle 1 or a new, unknown task, called Puzzle 2, which yielded more extreme payoffs than Puzzle 1.

At this time, participants were assigned to either the Baseline or one of the three treatments. In the treatments, pairs were formed according to the participant's and her partner's trust score with respect to the sample median. In the High-High treatment, players were matched with a partner with a similar and high trust score and were informed thereof. In the Low-Low treatment, players were matched with a partner with a similar and low trust score and were

informed thereof. In the Mixed treatment, players were matched with a partner with a different level of trust and were informed thereof. In the Baseline, participants were not matched with any partner. Moreover, in the three treatments each participant's payoff in the chosen task depended not only on her performance, but also on the partner's choice, creating a strategic interaction. In the Baseline, no strategic interaction took place. Here is a summary of treatment conditions:

High-High	**Low-Low**	**Mixed**	**Baseline**
High-trust participant matched with high-trust partner	Low-trust participant matched with low-trust partner	Mixed pairs	No matching
Information about the partner's trust	Information about the partner's trust	Information about the partner's trust	No info
Strategic interaction	Strategic interaction	Strategic interaction	No strategic interaction

Table 3: Treatments and Baseline in the experiment

The experiment ended with a third stage consisting of demographic questions on age and gender. Due to issues in participants' enrolment to sessions, we could not distribute them evenly across treatments. Instead, we assigned 44 subjects to the Baseline, 38 to "High-High", 40 to "Low-Low", 42 to "Mixed". Below we summarize the experimental flow in the Baseline and in the three treatments:

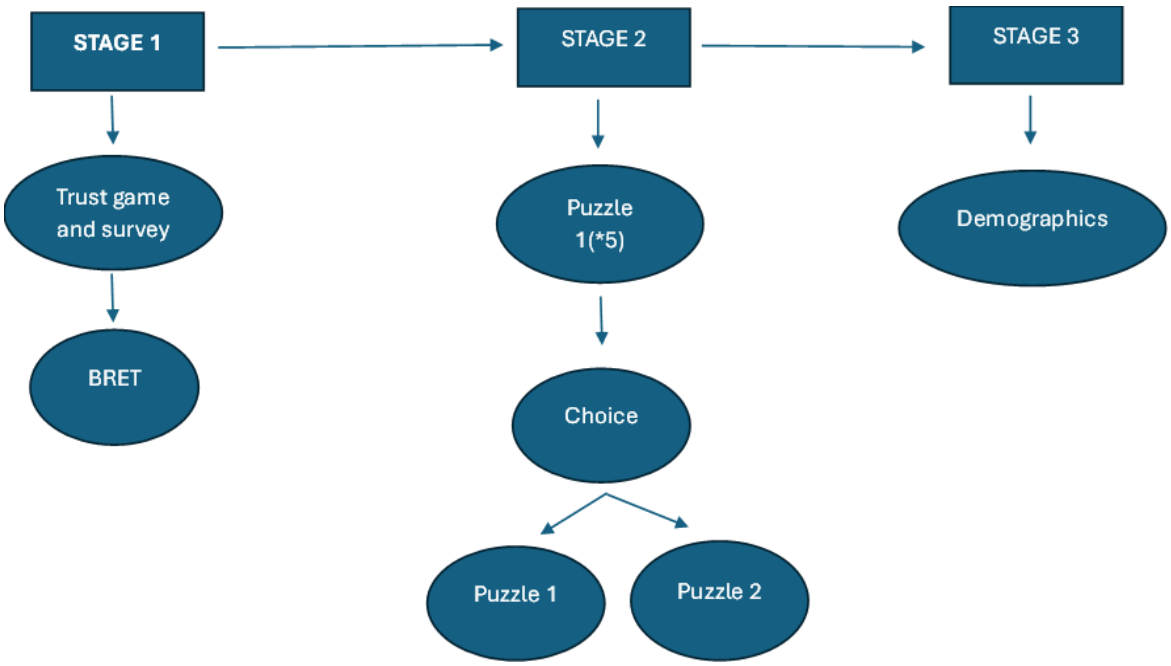


Figure 9: The experimental flow in the Baseline

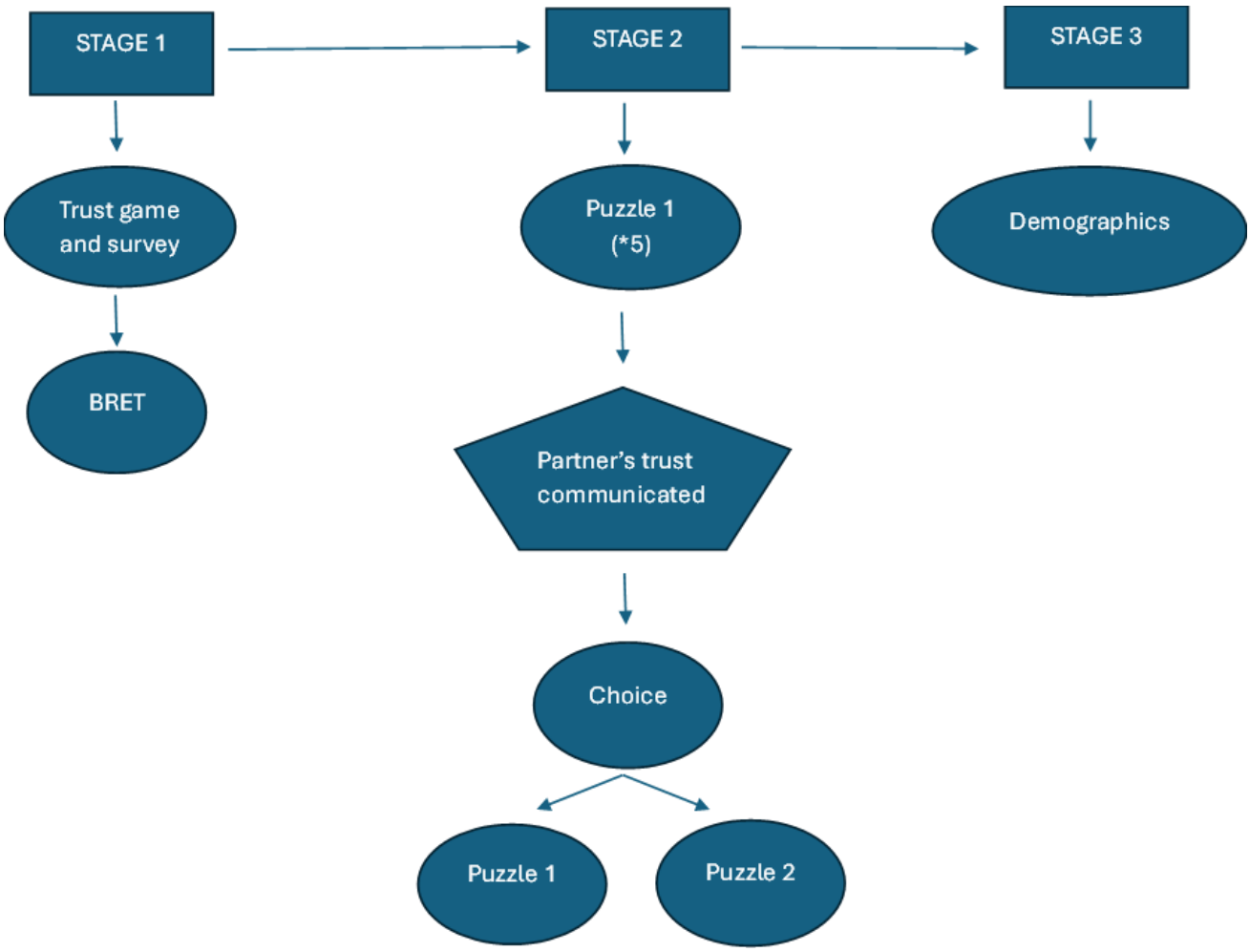


Figure 10: The experimental flow in the treatments High-High, Low-Low and Mixed

3.4.1 First stage

In the first stage, we elicited subjects' trust and risk attitude. Trust was measured using both a survey and an experimental measure. We used respectively the Rosenberg question from the World Values Survey (Ospina, Roser and Arriagada, 2016) and the trust game (Berg, Dickhaut and McCabe, 1995). We decided to combine the two measures to have a more precise estimate of trust and to compensate for the limitations inherent in each approach.

The Rosenberg question asks participants the following: "*Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people*?". Possible answers are in a discrete scale from 0, that is, "you can't be too careful" and 10, that is "most people can be trusted". The question was translated into Italian. Following the OECD guidelines on measuring trust (2017), response options were not dichotomous. This allows us to represent the whole range of scale values of trust, from low levels to medium and high levels. First ideated by Rosenberg in 1957, this measure is among the most widely used in trust surveys, including, besides the World Values Survey, the U.S. General Social Survey (GSS), the European Social Survey (ESS), the Civic Culture surveys, the American National Election Studies and the OECD database Trustlab (Murtin et al, 2018). In the latter case, the Rosenberg question is used in combination with experimental measures of trust, including the trust game. Its usage in many different contexts proves a high level of external validity. As reported by the OECD (2017), the Rosenberg question shows high levels of convergent validity, as it correlates with many indicators of well-being, including GDP per capita (e.g.,

Knack and Keefer, 1997), life expectancy, life satisfaction and lower unemployment (European Social Survey, 2002-2014). As trust is expected to have a positive effect on these variables, we can state that the Rosenberg question is good at measuring the concept it was designed to evaluate, thus having a good construct validity. As we explained in the introduction, Li, Turmunkh and Wakker (2019) show that survey measures of trust do not suffer from noise originated by ambiguity and strategic uncertainty, which characterize the trust game, and manage to exclusively capture beliefs about others' trustworthiness.

Moreover, we used the trust game (Berg, Dickhaut and McCabe, 1995) to add an experimental measure of trust. In fact, survey questions are subject to the risk of *social desirability bias,* that is, the likelihood that respondents overstate their level of trust to please the experimenter. Although this does not disqualify the Rosenberg Question as a measure of trust, we decided to add a second measure to obtain a more precise estimate. The trust game is a sequential game played in pairs, whereby one participant is player A, and her partner is player B. They are both endowed with 10 ECU (Experimental Currency Units). Player A moves first and decides how much of her endowment she wants to transfer to player B. The amount transferred is tripled, then player B decides how much to return to player B, knowing that she cannot return more than three times the sum transferred by player A. If we label the amount transferred by A and the amount returned by B respectively as t and r, then A's payoff is $10 - t + r$, and B's payoff is $10 + 3t - r$, with $r \leq 3t$. The amount transferred by A is our experimental measure of trust, while the amount returned by B is a measure of trustworthiness. Therefore, our final measure of trust was the average of each participant's answer to the Rosenberg question and the amount transferred in the trust game. As responses to the Rosenberg question tend to be skewed towards higher values and transfers in the trust game tend to have spikes around the middle value (5) and the maximum value (10), we believe that the two measures compensate themselves and return a more precise estimate of individual social trust.

In our experiment, the game was one shot. In the game, which was presented to them as Task 1, participants played both as player A and as player B. After receiving the instructions of the task, participants were asked first how much they wished to transfer to player B when they were player A, then how much they would return to each possible amount transferred by A when they were player B. In total, each player made 11 choices. This procedure is called strategy method (Selten, 1967), and it is widely used in the literature (e.g., Sapienza, Toldra-Simats and Zingales, 2013; Murtin et al., 2018). When playing as player B, participants were asked to fill in the third column of the same table that we used in the second chapter, knowing that they could not return more than the corresponding amount in the second column.

As in our second chapter, to calculate payoffs from the trust game half of the players were randomly assigned to the role of player A, the other half were assigned to the role of player B. Pairs were then formed, and for each pair the amount t transferred by player A was matched with the corresponding amount r returned by player B to calculate respective payoffs. Participants were informed of this procedure. Pairs were not the same as the ones created in the second stage. However, values of trust and trustworthiness were saved for all players for the data analysis, irrespective of their assignment to the role of player A or player B. The trust game was incentivised, and players could earn a maximum of 40 ECU, which corresponds to € 0.4. After the trust game, participants were asked to answer the Rosenberg question.

If Player A sent to you	You received	You send to A
1	3	
2	6	
3	9	
4	12	
5	15	
6	18	
7	21	
8	24	
9	27	
10	30	

Table 4: Choices of player B with the strategy method

We later proceeded to elicit risk attitude. We used the Bomb Risk Elicitation Task (BRET), originally developed by Crosetto and Filippin (2003). In this task, participants were shown a screen with 100 boxes. Each box was worth 1 ECU, but one of them contained a "bomb". Subjects had to choose how many boxes to collect, knowing that each box had the same probability of containing the bomb. If they avoided the bomb, their payoff equaled the number of boxes collected. However, if they collected the bomb, their payoff would be 0. As collecting more boxes increases the risk of selecting the bomb, the number of boxes collected is the experimental measure of risk attitude. The maximum payoff participants could gain from the BRET was 99 ECU, that is, € 0.99. As highlighted in the second chapter, the BRET offers many advantages compared to other measures of risk attitude. For example, the design is simple and intuitive, the task is entirely defined in the gain domain and noise generated by loss aversion is ruled out (Crosetto and Filippin, 2003). Unlike the second chapter, both the trust game and the BRET were incentivized. Below is the screen that participants saw. They had to write the number of boxes they wanted to collect and then click on "Conferma":

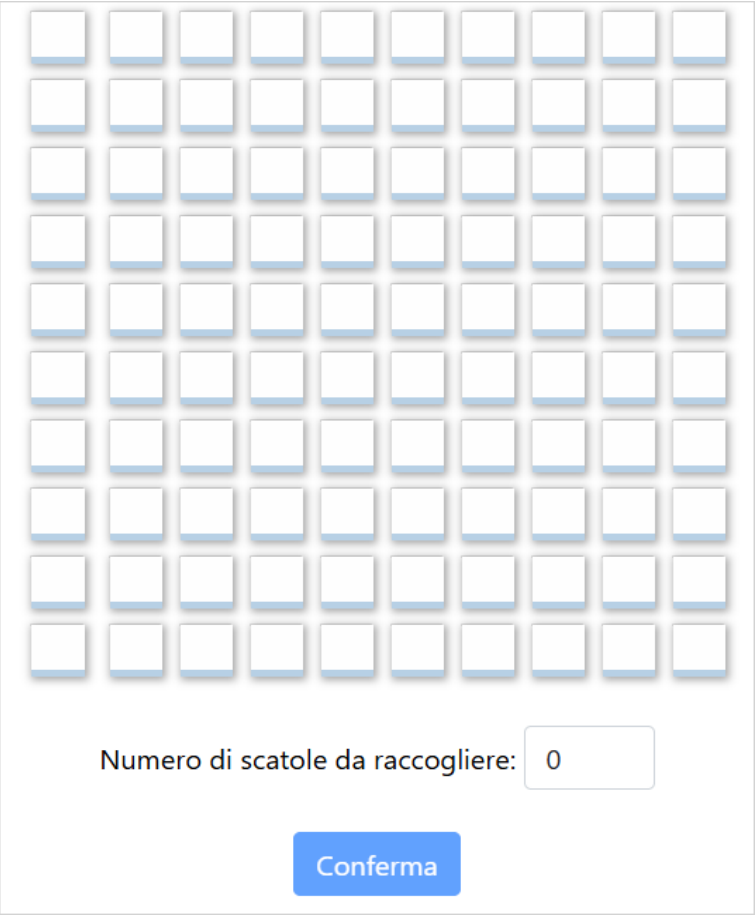


Figure 11: The BRET

3.4.2 Second stage

The second stage included a training phase and an active phase. In the training phase, participants familiarised themselves with a task, called Puzzle 1. By repeating the task for five rounds, participants could refine their choice through *learning by doing* and create a habit. With respect to the second chapter, we increased the number of rounds of the task in the training phase to reinforce the *lock-in* effect, whereby subjects develop a comfortable habit and are wary of any kind of innovation (Foxon, 2014). In the active phase, participants could choose whether to play another round of Puzzle 1 or to attempt the riskier, but potentially more rewarding Puzzle 2. Participants were informed that they would play six rounds of a task called Task 3, and that in the sixth round they would be given the opportunity to choose between two modes of play, called Puzzle 1 and Puzzle 2. This choice recalls the exploration-exploitation dilemma (March,1991), whereby organizations must choose between refining and improving a known technology or process or attempting to follow a new path. As in our setting, exploitation is safer than exploration, but the latter can potentially yield higher returns. Puzzle 1 and Puzzle 2 are real-effort tasks, where participants work to achieve outcomes (Carpenter and Huet-Vaughn, 2019). We chose real-effort tasks as they can reproduce situations happening in real life, inducing feelings such as fear, anxiety, boredom, curiosity and excitement.

Before making the choice, participants were paired according to their trust score in the first stage. In each session, trust scores were ordered in ascending fashion, and the median score was calculated. In the High–High and Low–Low treatments, participants were matched

sequentially: the lowest score with the second lowest, the third with the fourth, and so on. Those scoring below the median were assigned to the Low–Low treatment, while those above the median were assigned to the High–High treatment. In the "Mixed" treatment, instead, scores were divided into two groups around the sample median. Participants from the lower group were paired with participants from the upper group in matching order, that is, the first with the first, the second with the second etc. Before playing the chosen task, participants were also asked how they expected to fare in it. In the treatments, after playing either Puzzle 1 or Puzzle 2, participants were asked which task they believed the partner had chosen. The experiment ended with demographic questions on age and gender.

3.4.2.1 Puzzle 1

The training phase of the second stage consisted of five rounds of a puzzle building task, modelled after a similar task in Mittone and Papi (2017). In each round, the subject could see two grids on a screen. The right-side grid consisted of a target figure composed of black and white pixels, while the left-side grid was blank, made of white pixels. The task was to reproduce the target figure in the blank left-side grid by clicking on white pixels, so that they turned black. The number of moves was unlimited, and one could click on the same pixel infinite times. Each round had a time constraint of 30 seconds and a different target figure. Participants played the task alone.

An example of the task is provided in the figure below:

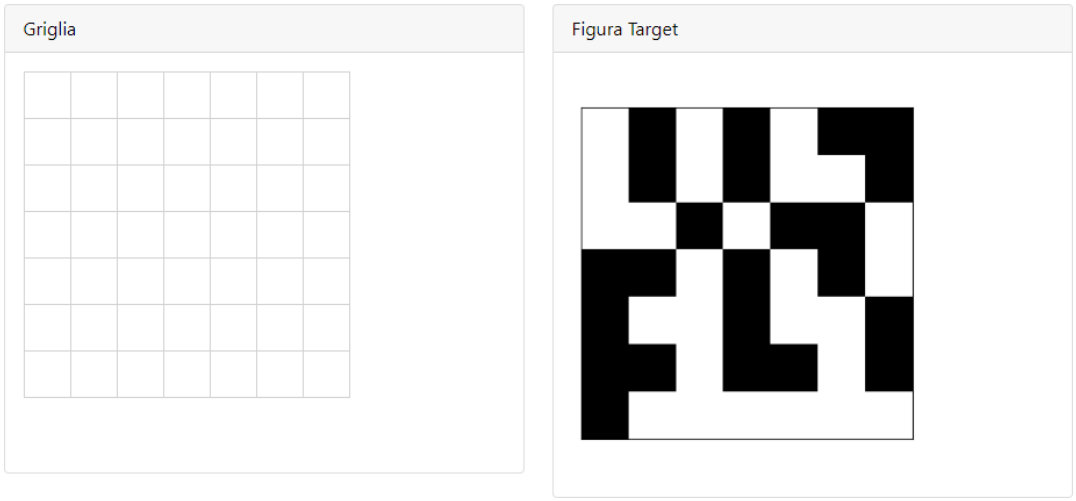


Figure 12: The Puzzle Task

Participants were told that their potential payoff depended on the number of mistakes they committed. In fact, in each round players gained 150 ECU if they reproduced the figure perfectly, 70 ECU otherwise. As the figures were relatively small and simple to reproduce, we decided to adopt a dichotomous payoff structure, which did not allow mistakes. Participants were then informed that at the end of the experiment one round of the puzzle task would be randomly chosen for payment.

3.4.2.2 Puzzle 2

After the training phase, participants could choose between playing another round of Puzzle 1 or attempting a new task, which we called Puzzle 2. In the three treatments, this active phase

involved a strategic interaction, whereby participants' payoff depended on the partner's choice. We will describe the details of the choice in the following subsection.

Puzzle 2 resembled Puzzle 1, but it was both more challenging and potentially more rewarding. As before, participants worked with two grids, but the solution strategy was radically different. On the right side of the screen, a target figure was displayed. On the left side, instead of a blank grid, participants saw a completed figure composed of black and white pixels, which did not match the target figure. The task consisted of transforming the left grid so that it reproduced the target figure. This required moving black pixels, one at a time, to the correct position. If participants clicked on a white pixel, nothing happened, while if they clicked on a black pixel, it turned grey, to show that it had been selected. The chosen pixel could then be moved to one of the four adjacent white squares in the grid. If a pixel was located on the edge of the grid, the number of adjacent available pixels was lower than four. Moves into spaces already occupied by black pixels were not permitted.

We provide a figure to illustrate the solution method of the task, where on the right side we have the target figure, and on the left side we have the grid which players could modify:

Figure 13: Example of Puzzle 2

The task was one-shot, with a time limit of 50 seconds. As for Puzzle 2, it had a dichotomous payoff, yielding 200 ECU if the figure was perfectly reproduced, 0 otherwise. Participants played the task alone. In the three treatments, a bonus or a malus could be added to the task's payoff according to the choice of the partner. Participants in the treatment were informed of their possible gains in Puzzle 2 depending on their performance and the two potential choices of the partner. We will explain the payoff structure in the following subsection.

3.4.2.3 Choice and strategic interaction

The choice was presented to subjects as one between two modes of play of Task 3. Puzzle 1 was the described as the mode they had played before, while Puzzle 2 was described as an unknown alternative mode, which was more challenging and could lead to larger gains in case of success and smaller gains in case of failure.

Before making her choice between Puzzle 1 and Puzzle 2, each participant was assigned to either the Baseline or one of the three treatments. As we explained in the previous subsection, in the three treatments each participant was matched with a partner according to the respective trust scores. Participants in treatments High-High, Low-Low and Mixed received the following information: "*We would like to inform you that you have been paired with another participant. You are among the half of participants with the highest (lowest) trust value, calculated on the basis of your choices in Task 1. The other participant is similar to (different from) you. Your*

earnings will depend on both your choice and that of the other participant.".

In this experiment, ambiguity and uncertainty originated from the unknown probability about what the partner would choose, the identity of the partner and, for participants who chose Puzzle 2, the nature of the task. Uncertainty regarding the type of the partner was mitigated by providing information about her level of trust and similarity with the participant.

Although participants played the task alone, in the treatments their payoff was affected by the partner's choice. A bonus or malus could thus be added to their payoff in the chosen task.

The payoff structure will thus include the following:

- If both players chose Puzzle 2, they both got a bonus of 30 ECU. The bonus was a prize for successful coordination on innovation.
- If one player chose Puzzle 1 and the partner chose Puzzle 2, the former got a bonus of 100 ECU, the latter got a malus of 70 ECU. Defecting means free-riding and yields a benefit, while innovating alone is more difficult and results in a loss. In fact, in our setting innovation in a dyad could only be successful if both members chose to innovate.
- If both players choose Puzzle 1, no bonus or malus was added.

We summarise the payoff scheme for each possible combination of choices of the two dyad members in the following table, where $\pi_1$is the row player's payoff, $\pi_2$is the column player's payoff, $p1$ is Puzzle 1 and $p2$ is Puzzle 2:

	Puzzle 1	**Puzzle 2**
Puzzle 1	$\pi_1(p1)$; $\pi_2(p1)$	$\pi_1(p1) + 100$; $\pi_2(p2) - 70$
Puzzle 2	$\pi_1(p2) - 70$; $\pi_2\ (p1) + 100$	$\pi_1(p2) + 30$; $\pi_2\ (p2) + 30$

This payoff structure creates a prisoner's dilemma. The Pareto-efficient outcome in which no participant can increase her well-being without harming the partner, is reached when both partners decide to innovate, that is, to choose Puzzle 2. This would maximize the dyad's payoff. However, unilaterally deviating and choosing Puzzle 1 always increases individual payoff, which is maximized if a participant chooses Puzzle 1 and succeeds, and her partner chooses Puzzle 2. Therefore, the only Nash equilibrium of the game is obtained when both players defect and choose Puzzle 1. Participants were not informed that they would play a prisoner's dilemma, instead they only knew that their payoff depended on the partner's choice.

The extensive-form game, which returns all possible payoffs of player 1 and player 2 for each combination of strategies and performance of both dyad members, can be represented by a game tree. Player 1 makes her choice between Puzzle 1 and Puzzle 2, and she can either fail or succeed in the task. Simultaneously, player 2 makes her choice and she can either fail or succeed in the task. Success or failure in the task happen with unknown probability, for example *p* for success in Puzzle 1 and *q* for success in Puzzle 2. Probabilities of success are different for each subject. Participants were never informed of whether the partner had succeeded in her task, but at the end of the experiment they got to know what task the partner had chosen. When both players have chosen and played their task, either succeeding or failing, their payoffs are indicated in parentheses, with the first number representing player 1's payoff and the second number representing player 2's payoff. For example, (-70, 250) is the pair of payoffs corresponding to a situation where player 1 played Puzzle 2 and failed, while player 2 played Puzzle 1 and succeeded. Here is the resulting game tree:

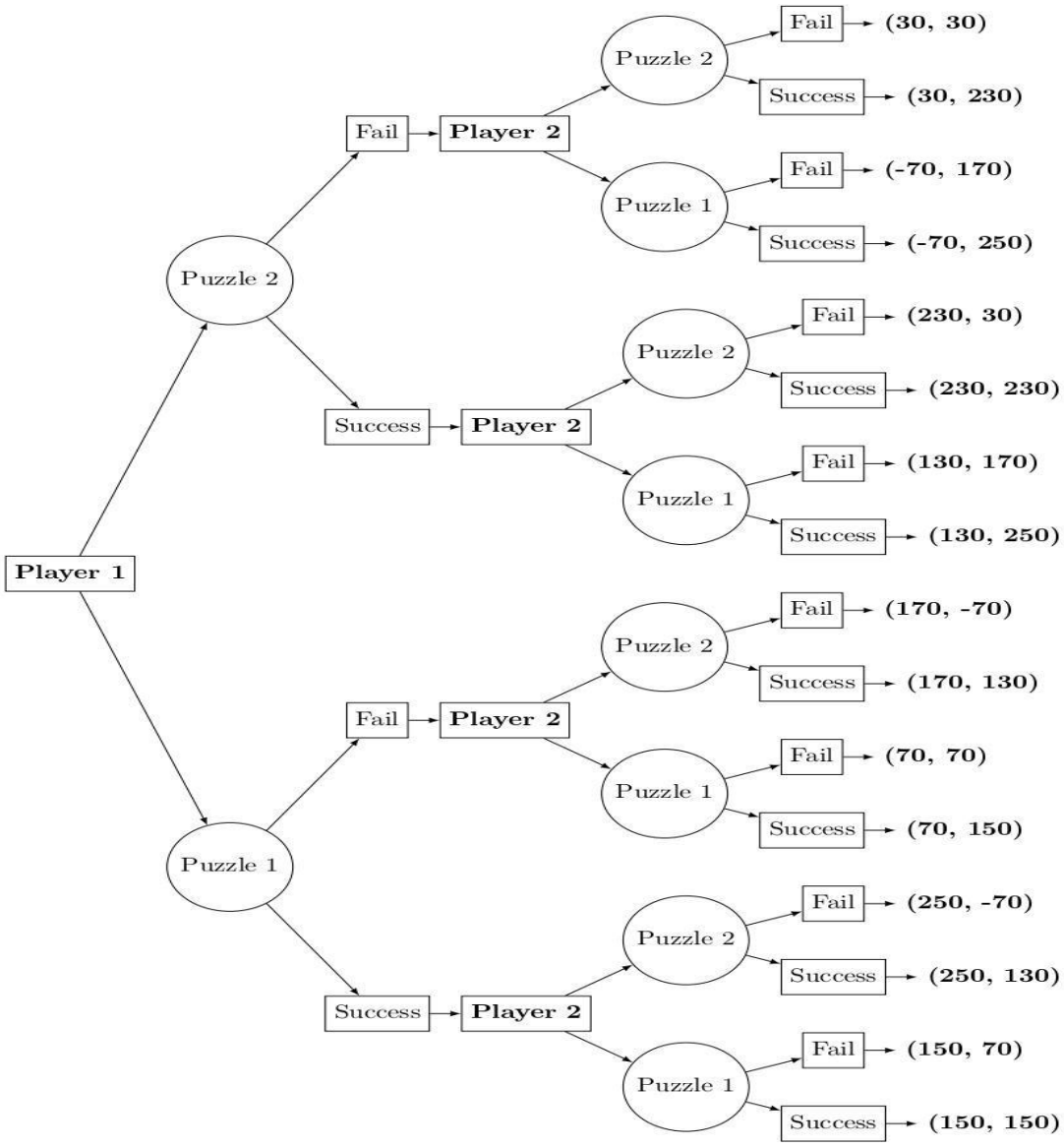


Figure 14: The complete game tree

Before choosing, players were shown a simplified and interactive version of this game tree, where they could see their payoff and the partner's payoff for each possible combination of choices and performances. We chose to show participants this simplified decision tree to avoid ceiling effects favoring one of the two tasks, and to ensure they were aware of the consequences of their choices for themselves and the partner. Players could thus understand that defecting was the dominant strategy, but they could also be induced to trust that their partner would choose to follow them in innovation and maximize the dyad's payoff. Here is an example of the tree that participants saw before choosing. They had to select all possible combinations of strategies and to check the possible payoff for them and the partner before being allowed to proceed to the choice.

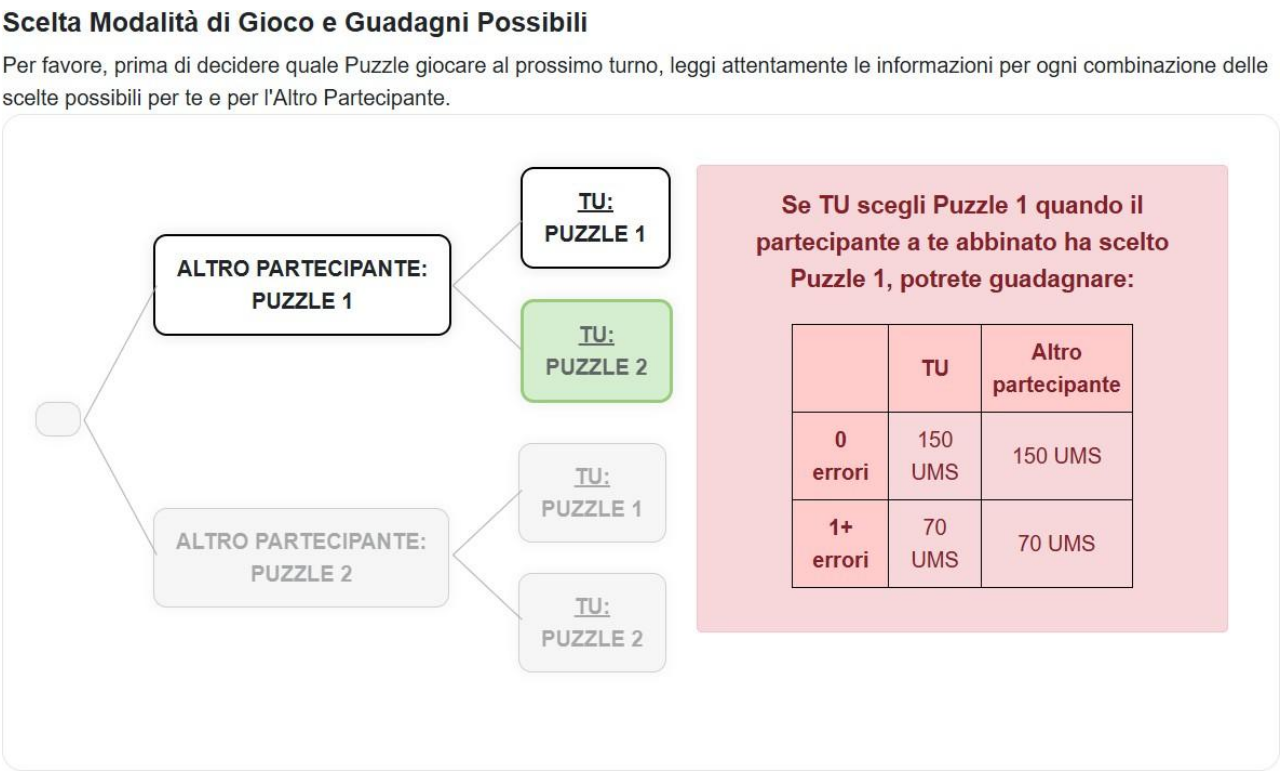


Figure 15: How the game tree was shown to participants in the treatments

As participants made their choice, they were asked to rate their probability of success in the chosen task, from 1 to 5. Participants in the Baseline were asked the same question. After playing the chosen task, each participant in the treatments was asked what task she believed the partner had chosen. This allowed us to control for self-perceived probability of success and beliefs about the partner's choice. We believe that these variables could play a significant role in participants' choice.

3.4.3 Demographics

The experiment ended with demographic questions on each participant's nationality and gender. When asked about their gender, participants had the possibility to reply "non-binary" or "Prefer not to declare".

3.5 Results

In this section we describe the results of our experiment. To conduct our data analysis, we ran a series of logistic regressions on R. We decided to use this type of regression because the

dependent variable is either *Willingness to Innovate* or *Belief about the Partner*, which are both dummy variables. The former is equal to 1 if the participant chose Puzzle 2, 0 otherwise. The latter is equal to 1 if the participant thought that her partner chose Puzzle 2, 0 otherwise.

The main explanatory variables are *Trust*, that is, the average between the amount transferred in the trust game and the answer to the Rosenberg Question, and *Partner's Trust,* a dummy equal to 1 if the partner's trust score was higher than the sample median, 0 otherwise. Dummies for the treatments and the Baseline were created, and we introduced interaction variables between the aforementioned dummies and Trust. We also created the dummy *Partner's Trust*Trust* to control for a moderating effect of Trust on the relationship between Partner's Trust and Willingness to Innovate.

3.5.1 Hypothesis 1

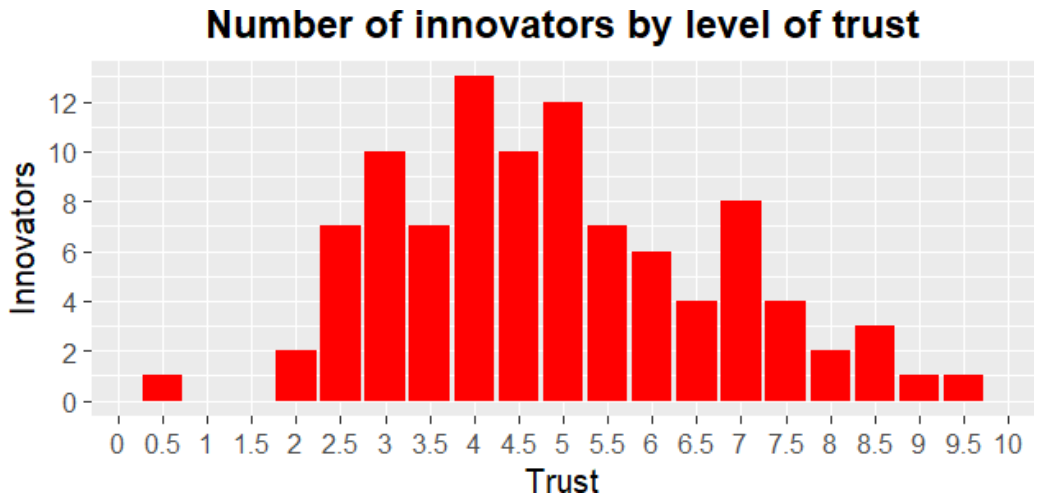
To test for Hypothesis 1, we ran a logistic regression, whose dependent variable was *Willingness to Innovate* and whose dependent variable was *Trust*. As expected, the coefficient of Trust is not significant, as the observations include participants from the Baseline, who did not engage in a strategic interaction. In fact, trust cannot be a driver of innovation when no strategic interaction happens.

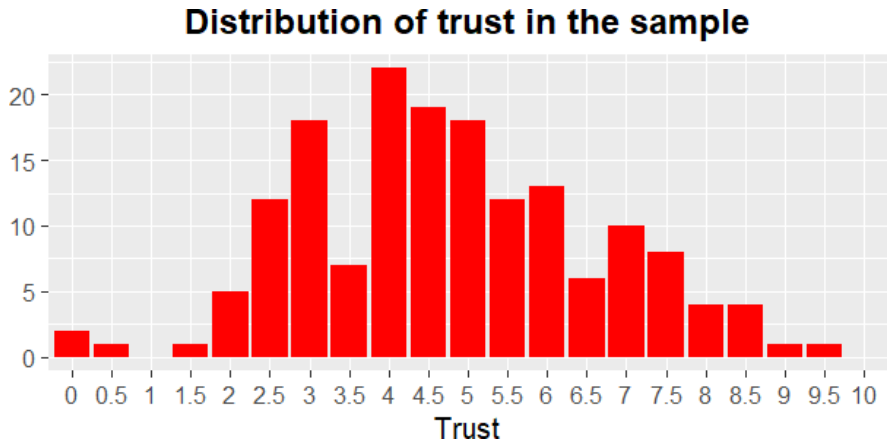
Trust and Innovation

	Dependent variable:
	Willingness to Innovate
Trust	0.100 (0.088)
Constant	-0.076 (0.445)
Observations	164
Log Likelihood	-109.889
Akaike Inf. Crit.	223.779

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

We represented the absolute frequency of innovators, that is, those who chose Puzzle 2, by level of Trust with a bar plot. Levels of Trust include decimal numbers as they are an average between scores in the trust game and answers to the Rosenberg question. Results do not show any clear trend, except a spike around levels from 4 to 5. We compared this bar plot to the distribution of Trust across the sample, which seems to show a relatively similar trend:





If we include controls, we see that *Belief about the Partner* has a positive and significant effect on Willingness to Innovate. This means that subjects understood the strategic nature of the choice and mostly chose to innovate when they expected the partner to do the same. The coefficient for Trust remains non-significant, and the same holds for other explanatory variables. 44 observations were lost because the variable Belief about the Partner is not defined in the Baseline.

Trust and Willingness to Innovate with controls

	Dependent variable:
	Willingness to Innovate
Average trustworthiness	0.047 (0.057)
Trust	0.182 (0.124)
Risk attitude	-0.002 (0.010)
Male	1.386 (1.299)

Female	0.729 (1.308)
Age	-0.051 (0.053)
Belief about the Partner	0.748* (0.402)
Belief about self	-0.382 (0.249)
Puzzle task ability	0.033 (0.140)
Constant	0.329 (2.274)
Observations	120
Log Likelihood	-75.774
Akaike Inf. Crit.	171.548

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

To conclude our analysis on Trust and Willingness to Innovate, we investigated whether the effect of Trust was different across the treatment conditions. The first step was to create the dummy *Treatment*, which is 0 if the participant was assigned to the Baseline, 1 otherwise. This is the result of the regression:

Trust and assignment to treatment

	Dependent variable:
	Willingness to Innovate
Trust	-0.120 (0.188)
Treatment	-2.009* (1.119)
Trust*Treatment	0.287 (0.215)
Constant	1.454 (0.991)
Observations	164
Log Likelihood	-107.513
Akaike Inf. Crit.	223.026

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

Assignment to a treatment condition other than the Baseline has a negative and significant effect on Willingness to Innovate. This was expected, because the treatments include a strategic interaction with a partner, which entails the risk of unilateral defection and betrayal and the presence of ambiguity and strategic uncertainty.

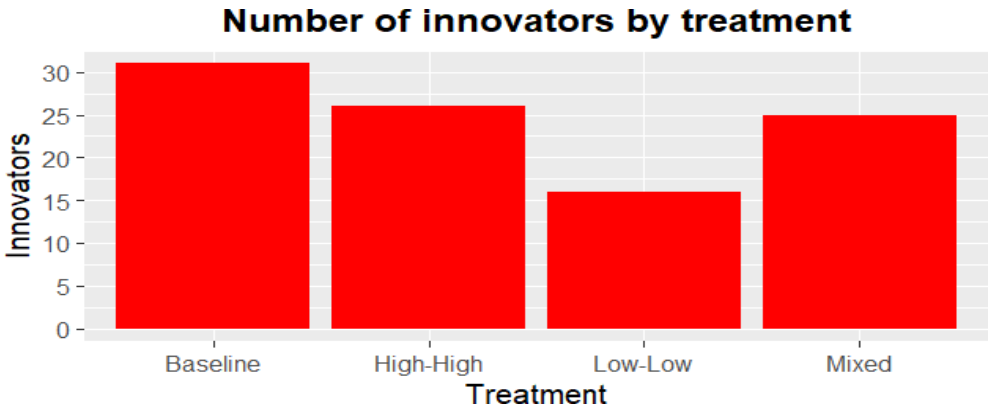
To conclude, we studied the effect of Trust on Willingness to Innovate across the three treatments. We created the dummy variables *High-High*, *Low-Low* and *Mixed,* and the interaction variables between these dummies and Trust. The dummy *Baseline* was used as a reference point to avoid multicollinearity. No evidence of significant differences across treatments is provided by the data.

Trust and willingness to innovate across treatments

	Dependent variable:
	Willingness to Innovate
Trust	-0.120 (0.188)
High-High	-2.731 (1.901)
Low-Low	-1.860 (1.474)
Mixed	-0.717 (1.301)
Trust*High-High	0.470 (0.335)
Trust*Low-Low	0.120 (0.352)
trust*Mixed	0.047 (0.247)
Constant	1.454 (0.991)
Observations	164
Log Likelihood	-104.484
Akaike Inf. Crit.	224.968

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

Here is the bar plot comparing the frequency of innovators in the Baseline and in the three treatments:



We see that the number of innovators is the highest in the Baseline and the lowest in the Low-Low treatment. As the number of innovators is the lowest in a treatment when participants are exclusively matched with low-trust partners and it is notably higher in treatments when there is at least the possibility to be matched with a high-trust partner, we could suppose a possible effect of Partner's Trust on Willingness to Innovate. However, this will be investigated in the next subsection.

To conclude, the data analysis does not provide us with any evidence of correlation between Trust and Willingness to Innovate. Our Hypothesis 1 cannot be confirmed.

3.5.2 Hypothesis 2

To test for Hypothesis 2, we filtered the sample for participants belonging to the three treatments, excluding those belonging to the Baseline. We ran a logistic regression whose dependent variable was *Willingness to Innovate*, and whose explanatory variables were *Trust* and *Partner's Trust*:

$$Willingness\ to\ Innovate_i = \alpha + \beta Trust_i + \gamma Partner's\ Trust_i + \varepsilon_i$$

Results show a strong positive correlation between Partner's Trust and Willingness to Innovate, demonstrating that participants who were matched with a high-trust partner were more likely to choose to innovate.

Partner's Trust and Willingness to Innovate

	Dependent variable:
	Willingness to Innovate
Trust	0.142 (0.106)
Partner's Trust	0.770** (0.381)
Constant	-0.802 (0.542)
Observations	120
Log Likelihood	-78.939
Akaike Inf. Crit.	163.878

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

If we include controls, the coefficient for *Partner's Trust* remains positive and significant. The same holds for the coefficient for Belief about the Partner.

Partner's trust with controls

	Dependent variable:
	Willingness to Innovate

Average Trustworthiness	0.050 (0.058)
Trust	0.160 (0.127)
Partner's Trust	0.774* (0.405)
Risk attitude	-0.001 (0.010)
Age	-0.053 (0.057)
Female	0.878 (1.340)
Male	1.608 (1.341)
Belief about self	-0.381 (0.254)
Belief about the partner	0.675* (0.409)
Puzzle task ability	0.050 (0.144)
Constant	-0.119 (2.322)
Observations	120
Log Likelihood	-73.912
Akaike Inf. Crit.	169.824

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

Given the results of the last regression, we decided to investigate the relationship between Belief about the Partner and Partner's Trust. Results show a positive correlation, suggesting that knowing that the partner is a high-trust type could lead participants to believe that the partner had decided to innovate. Information that the partner is a high-trust type seems to generate feelings of expected trustworthiness and expected reciprocity regarding the partner, thus leading to a higher degree of cooperation. This is in line with the literature (e.g., Yamagishi and Kiyonari, 2000). However, if we include controls, the coefficient of Partner's Trust loses significance, thus suggesting that the effect of Partner's Trust on Willingness to Innovate is not completely mediated by Belief about the Partner.

Beliefs about the partner, Trust and Partner's Trust

	Dependent variable:
	Belief about the partner
Trust	0.024 (0.025)
Partner's Trust	0.153* (0.092)

Constant	0.284** (0.129)
Observations	120
Log Likelihood	-85.781
Akaike Inf. Crit.	177.562

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

Belief about the partner and controls

	Dependent variable:
	Belief about the partner
Average trustworthiness	-0.008 (0.013)
Trust	0.040 (0.028)
Partner's Trust	0.138 (0.093)
Risk attitude	-0.001 (0.002)
Age	-0.020* (0.012)
Female	-0.237 (0.302)
Male	-0.191 (0.299)
Belief about self	-0.090 (0.056)
Puzzle task ability	0.036 (0.031)
Constant	1.303** (0.515)
Observations	120
Log Likelihood	-82.258
Akaike Inf. Crit.	184.517

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

To conclude, our results allow us not to reject Hypothesis 2. Informing subjects that the partner is a high-trust type seems to foster Willingness to Innovate. A possible channel could be beliefs about the partner's choice, but other factors may also come into play.

3.5.3 Hypothesis 3

To test for Hypothesis 3, we ran a logistic regression whose dependent variable was Willingness to Innovate and whose explanatory variables were Trust, Partner's Trust and the interaction variable *Trust*Partner's Trust.* The regression is the following:

$$Willingness\ to\ Innovate_i = \alpha + \beta Trust_i + \gamma Partner's\ Trust_i + \ \delta Trust_i * Partner's\ Trust_i + \varepsilon_i$$

The effect of Partner's Trust on Willingness to Innovate is thus equal to $\gamma + \delta Trust_i$. The sign of the coefficient δ determines the direction of the moderation effect. A negative δ would suggest that the effect of Partner's Trust is stronger for low-trust individuals, thus confirming our hypothesis. However, our results do not support Hypothesis 3. Although the coefficient δ is positive, it is not statistically significant. This lack of significance does not allow us to draw any conclusion about whether the effect of knowing that the partner is a high-trust type differs between subjects with high or low individual trust.

Interaction between Trust and Partner's Trust

	Dependent variable:
	Willingness to Innovate
Trust	0.064 (0.141)
Partner's trust	-0.047 (1.076)
Trust*Partner's Trust	0.174 (0.216)
Constant	-0.454 (0.687)
Observations	120
Log Likelihood	-78.608
Akaike Inf. Crit.	165.215

*Logistic regression estimates. Standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01*

3.6 Discussion

Our third chapter aimed at investigating the relationship between trust, partner's trust and willingness to innovate in an experimental setting. We hypothesized that trust and partner's trust would have fostered willingness to innovate, and the effect of partner's trust would have been stronger for subjects with low individual trust.

We found evidence supporting a correlation between partner's trust and willingness to innovate, that is, participants were more willing to innovate when they knew that they had been matched with a high-trust partner. This confirms our Hypothesis 2, and it is the most relevant finding of our work. Information on the partner's personality traits appears to be relevant when engaging in a strategic interaction. This is especially true when the task is one-

shot and no information on the partner's previous behavior is available (e.g., Guilfoos and Kurtz, 2017). In our setting, innovation leads to higher payoffs for the dyad only when both members manage to coordinate on choosing the new task. As we have seen in the second chapter, innovation in modern times mostly happens within large firms or network of firms, which have teams dedicated to research and development (e.g., Powell, Koput and Smith-Doerr, 1996). In these teams, the decision to innovate must be taken collectively. Members need to coordinate on innovation; therefore, innovators must have the courage to communicate their ideas, and they need to trust that other members will follow them. Knowledge of the partners' personality traits can help innovators successfully communicate their ideas and win the colleagues' support. In our thesis, we found evidence that the effect of information on the partner's trust on willingness to innovate is at least partially mediated by beliefs about the partner's choice. This is consistent with the literature, according to which most players cooperate in social dilemmas only when they expect to be reciprocated (e.g., Rabin, 1993; Hayashi et al., 1999). If this is true, participants could misinterpret the prisoner's dilemma as an assurance game or a similar social exchange context, whereby cooperation is the optimal strategy when the partner is likely to reciprocate (e.g., Hayashi, 2000; Kiyonari, Tanida and Yamagishi, 2000). Conditional cooperation based on expected reciprocity has been seen by some authors as a successful strategy acquired throughout human evolution (Cosmides and Tooby, 1989; Gintis, 2002). Our manipulation does not directly refer to the probability of reciprocity, instead it refers to trust, that is, a personal trait that participants could interpret as a signal of higher trustworthiness and thus expected reciprocity. This would stimulate the process of *projection* (Acevedo and Krueger, 2005), whereby players anticipate the probability that their action will be reciprocated and decide their strategy accordingly. Manipulating projection has been proven as a possibly successful strategy to induce cooperation. Moreover, information on the partner's trust could have helped reduce the ambiguity and strategic uncertainty that characterize strategic settings. In fact, since we did not measure ambiguity aversion, ambiguity-generated insensitivity and ambiguity-neutral first-order beliefs, information on the partner's trust might have changed subjective probabilities about the future choice of the partner, thus making the decision to innovate easier to take.

However, expected reciprocity is not the only possible explanation for cooperative behavior. In fact, the relationship between Belief about the Partner and Partner's Trust loses significance when controls are added. Instead, other factors may come into play. For example, participants could also interpret high Partner's Trust as a personality trait that simply increases the likelihood that she will play cooperatively, without thoughts about possible reciprocity. Participants could thus decide to *assimilate* (Acevedo and Krueger, 2005) to the partner's most likely action and act cooperatively. In fact, high partner's trust could be interpreted as a signal of high honesty and morality (Van Lange and Kuhlman, 1994), or of personality traits such as agreeableness or openness to experience. Participants could thus be induced to act cooperatively and innovate, either because they feel the need to adhere to social norms or due to guilt aversion (Battigalli and Dufwenberg, 2007). Social norms such as inequity aversion could be perceived as more cogent once participants receive a clue about the partner's honesty. This is especially true among prosocials, who interpret social dilemmas in terms of morality rather than strength. Participants may thus fear to be perceived as less moral than the partner if they decide to deviate from innovation, and this might lead even low-trust participants to switch to a cooperative strategy. Ahn et al. (2001) identify fear of

partner's unilateral defection as one of the main drivers of defection in a prisoner's dilemma. This is coherent with Neto, Felipe and Bento Caleiro (2019)'s finding that fear is one of the main obstacles to innovation within teams. In fact, team members might feel that their innovative ideas might be ridiculed or ignored, or that the firm would not share the benefits of innovation with them (Clegg et al., 2002). If partner's trust is seen as a proxy of morality and kindness, the fear of opportunistic behavior by the partner might be alleviated, thus making cooperation for innovation more likely.

In a nutshell, we can tell that receiving information on the partner's level of trust makes the choice to cooperate on innovation more likely in a dyad involved in a strategic interaction. Future research may shed light on the exact channel that drives this correlation. Expectations of reciprocity and beliefs of morality seem to be the most likely explanations judging by the literature that we know, but other findings may come for future experimental studies. New variables and tasks might be added to the experiment. It would be important to measure attitudes about ambiguity and strategic aversion, as well as ambiguity-generated insensitivity, ambiguity-neutral first-order beliefs and betrayal aversion. A measure of guilt aversion could also be included. Moreover, we could introduce measures of other personality traits, including Social Value Orientation (SVO) and the Big Five Personality traits. Another possible development could be to include treatment conditions that involve the manipulation of the *projection* and *assimilation* processes described by Acevedo and Krueger (2005). In fact, it might be interesting to understand whether participant choose their strategy based on the likelihood of reciprocation or on the partner's most likely action. This could help explain how beliefs on the partner's action are formed, and what role information on the partner's personality trait plays. However, it is extremely important not to impose an excessive cognitive load on experimental subjects. Our experiment already lasts about 30 minutes and involves three different tasks. Extensions are welcome, but the longer the experiment, the higher the risk that boredom effect influences players' choice and weakens the internal validity of the experiment.

As regards the relationship between trust and willingness to innovate, this experiment confirms the lack of significance that we found in the second chapter of the thesis. While in the second chapter we attributed this to the small sample size or to possible flaws in the experimental design, this chapter seems to suggest that it is not individual characteristics that influence willingness to innovate in a strategic setting. In fact, none of the individual traits that we measured in the experiment, namely average trustworthiness, risk attitude, puzzle task ability or belief about personal success in the chosen task is significant. A possible explanation might be that in strategic settings where dyads play a one-shot prisoner's dilemma, information on the partner might be more significant than individual traits. However, further research is needed to confirm what would be a noteworthy finding. We should not forget that variables such as ambiguity and betrayal aversion, strategic uncertainty and Social Value Orientation were not included in the sample. These individual traits might play a role in determining participants' willingness to innovate.

As in the second chapter, we found that assignment to a treatment condition involving a strategic interaction negatively affects willingness to innovate. In fact, in the Baseline participants do not have to worry about the risk, ambiguity and uncertainty of a strategic interaction. Their success in either Puzzle 1 or Puzzle 2 only depends on their performance, therefore they feel more at ease innovating. Moreover, we did not find any evidence of

differences in the effect of partner's trust on willingness to innovate between high-trust and low-trust individuals. The interaction coefficient between Trust and Partner's Trust is not significant; therefore, no moderating effect is suggested by our results. Future research might shed light on this point.

To conclude, we need to describe the limitations of our study. Firstly, the sample size could be widened to increase power, so that smaller effects can be detected. Secondly, our experiment was conducted with students from the University of Trento. This choice simplified sample selection and data collection, but it also made the experiment less realistic. To overcome this limitation, we suggest replicating the experiment with employees working in companies, even if this could increase the costs of the experiment. However, it would be interesting to study how real workers react to the experimental conditions. A third limitation of our study is that beliefs on other participants' choice were only elicited in the treatments. We made this choice because in the Baseline no strategic interaction happens. However, we believe that it may be interesting to observe whether participants not involved in a strategic interaction think that other participants had innovated, and whether this has some effect on their choice. To conclude, not eliciting participants' beliefs on others' choice caused a loss of sample size in the regression, as only 120 observations out of 164 were considered. This caused the power of our experiment to shrink, thus limiting our ability to detect correlations.

3.7 Conclusions

In the present paper, we investigated the relationship between trust, partner's trust and willingness to innovate in dyads involved in a strategic interaction. We conducted a two-stage experiment with students from the University of Trento. In the first stage, we elicited trust and risk attitude. The second stage involved a training phase, where players familiarised themselves with a task, called Puzzle 1, and an active phase, where players could choose between playing another round of Puzzle 1 or attempting a new task, Puzzle 2, which was riskier and potentially more rewarding. The choice was modelled after the exploration-exploitation dilemma (March,1991), with Puzzle 2 being the innovative task. We created a dummy called Willingness to Innovate, which was equal to 1 when the participant chose Puzzle 2. In the second stage, participants were assigned to either the Baseline or one of three treatments. In the treatments, participants were matched with a partner based on respective levels of trust. Moreover, they were informed of their partner's level of trust and their payoff in the chosen task depended on the partner's choice. A prisoner's dilemma setting was created, whereby attempting Puzzle 2 meant cooperation and continuing with Puzzle 1 meant defection.

Although mutual defection is the only Nash equilibrium in a prisoner's dilemma, we expected to find evidence of cooperation. Trust and partner's trust, according to our hypotheses, could have helped dyads escape the bad equilibrium. We hypothesised that trust and partner's trust would have boosted willingness to innovate, and the effect of partner's trust would have been stronger for subjects with low individual trust. Results only partially supported our hypotheses. While no correlation between trust and willingness to innovate was found, partner's trust appeared to be positively and significantly correlated with willingness to innovate. Receiving the information that the partner was a high-trust type encouraged participants to choose the new task. The effect was at least partially mediated by expectations on the partner's choice, suggesting that high Partner's Trust could have been seen as a proxy of high expected reciprocity, which usually leads to higher cooperation in social dilemmas.

However, other factors could have come into play. For example, participants could have thought that high-trust partners scored high in honesty or other personality traits linked with higher morality or trustworthiness. This could have activated social norms favoring cooperation or the feeling of guilt aversion. Further research could help shed light on the exact channel through which partner's trust affects willingness to innovate. We did not find evidence of a different effect of partner's trust on willingness to innovate between participants with high and low individual trust. As expected, participants assigned to treatments involving a strategic interaction were less likely to choose to innovate, but information about partner's trust partially helped reduce the fear of opportunistic behavior by the counterpart.

Our research adds to the literature on trust and innovation while bridging it with the literature on personal traits, beliefs and cooperation in social dilemmas. Trust is seen as a coordination issue within a prisoner's dilemma setting, and the psychological motives driving the choice to innovate are the focus of our study. The experimental method is a feature of novelty with respect to most studies about trust and innovation. New studies might apply the same research method to investigate the motives of innovation in teams. Measures of other personality traits or manipulations of beliefs about the partner might be added to the experimental design. Moreover, we believe that further research with a bigger sample size and participants recruited from firms could help overcome the limitations of our study.

4 Conclusive remarks

In this last section, we summarize the results of our thesis and propose some suggestions for future research. The aim of this thesis was to investigate the relationship between trust and innovation within teams deciding collectively about future strategies. More specifically, we focused on dyads which could decide between exploitation, that is, refining existing technologies and processes, or exploration, that is, searching for new solutions. The choice was experimentally modelled as one between a well-known task, which participants had played multiple times, or a new, unknown task. Coherent with the exploration-exploitation dilemma, in our experiments the well-known task yielded a safer payoff, while the new task yielded a higher payoff in case of success and a lower payoff in case of failure. Intending to recreate the internal conflicts in a Research and Development team which must coordinate on whether to innovate, we nested the choice between the choice between the two tasks in a prisoner's dilemma setting. In both the second and third chapter, the payoff structure of the second stage was such that coordinating on innovation guaranteed the highest joint payoff, but unilateral defection towards the well-known tasks always maximised individual payoff. Even if, for the sake of simplicity, members of the dyads did not play the task together, a strategic interaction occurred, whereby each participant's payoff depended on the partner's choice. In this context, we wanted to measure the effect of trust and, in the third chapter, partner's trust on participants' willingness to innovate. The latter variable was measured as a dummy equal to 1 if the participant had chosen the new task.

The literature, which we reviewed in the first chapter, mainly suggested a positive correlation between trust and innovation. Possible drivers of this correlation might have been easier knowledge sharing and conflict management, lower monitoring costs and reduced fear of ridicule or exploitation from the counterpart. The literature warned of the risk of a possible nonlinear effect of trust on innovation, with excessive levels of trust having a potentially negative effect on innovation. However, in our two experimental chapters we did not find any evidence of a correlation between trust and innovation. Instead, we found a positive and significant effect of partner's trust on willingness to innovate. In fact, information about the partner's trust might be interpreted as a signal of higher expected reciprocity or morality. Receiving these clues could make subjects more likely to cooperate on innovation in the prisoner's dilemma. In fact, we gained hints of a possible correlation between information about partner's trust and the belief that she had played the new task. When people received the information that the partner was a high-trust type, they were more likely to believe that she chose to cooperate on innovation. While the effect is not particularly strong, it could suggest that some participants see partner's trust as a proxy for higher expected reciprocity. Other channels might also drive the effect of information on the partner's trust on willingness to innovate, including the desire to adhere to social norms and the feeling of guilt aversion which may arise if participants perceive the partner as "moral".

Both in the second and third chapter we find that belief that the partner had chosen to innovate is positively correlated with individual willingness to innovate, suggesting that participants did understand the strategic nature of the choice. Future research could focus on how beliefs about the partner are formed and how these beliefs influence subjects' choice in a prisoner's dilemma. For example, receiving hints about the partner's score in the Social Value Orientation, the Big Five Personality Traits or other personal characteristics might help

subjects formulate a hypothesis on the partner's trustworthiness and probability of reciprocation (e.g., Van Lange and Kuhlman, 1994; Guilfoos and Kurtz, 2017).

In a social dilemma, subjects could either hypothesize the partner's probability of reciprocation and decide which action to take, a process called projection, or anticipate the partner's choice and decide accordingly, a process called assimilation (Acevedo and Krueger, 2005). It is of utmost importance to understand which of these mechanisms is more relevant for choices in strategic situations and whether manipulating assimilation or projection yields differences in subjects' choices. Moreover, future studies should not overlook the importance of ambiguity and strategic uncertainty, which characterize settings such as the prisoner's dilemma. In fact, in our setting ambiguity and uncertainty arise from the unknown probability of reciprocation, the identity and type of the partner and the nature of the new task. Information about the partner's choice partially reduces uncertainty related to the partner and her possible choices, and this could partially explain the positive effect on willingness to innovate. All the variables and methodologies that we described could be included in future research, but one must be wary not to exaggerate the cognitive load of the experiment on subjects.

As we have outlined throughout our thesis, experiments have many advantages, including an easier identification of causal effect and a high level of replicability, therefore we suggest that future research is conducted using experimental techniques. To overcome the limitations of our second and third chapters, bigger samples should be collected, possibly including employees from real firms. While the higher sample size increases the power of the experiment, choosing participants from real firms might improve the realism of the design. However, conducting an experiment with real-life employees will certainly be more costly and time-consuming.

To conclude, we believe that it is of utmost importance to study the relationship between trust, partner's trust and willingness to innovate within teams. In fact, many organisations struggle with innovation and rely excessively on exploitation rather than exploration (March,1991). While this strategy could yield safer profits in the short term, in the long term the firm could lose market opportunities and profitability. The lock-in effect hinders change in firms' strategies, technologies and processes, and firms rarely manage to escape from this vicious cycle. Therefore, it is important to find strategies to overcome this impasse, and create a trusting environment when employees not only have high trust in others, but they know that colleagues trust them could be a solution that managers and entrepreneurs should consider. The same is true for public organisations dealing with innovation, such as research centres or universities.

5 Bibliography

Acevedo, M. and Krueger, J.I. (2005). 'Evidential Reasoning in the Prisoner's Dilemma', *The American Journal of Psychology.* 118(3): 431-457

Aghion, P., and Howitt, P. (1992). 'A Model of Growth Through Creative Destruction'. *Econometrica.* 60 (2): 323-351

Ahmed, A.M. (2007). 'Group identity, social distance and intergroup bias'. *Journal of Economic Psychology.* 28 (3): 324-337

Ahn. T.K., et al. (2001). 'Cooperation in PD Games: Fear, Greed, and History of Play'. *Public Choice*. 106 (1/2): 137-155. Available at: https://www.jstor.org/stable/30026189?seq=1

Akçomak, S., and ter Weel, B. (2009). 'Social capital, innovation and growth: Evidence from Europe'. *European Economic Review.* 53 (5): 544-567

Al-Ubaydli, O., Jones, G., and Weel, J. (2015). 'Average player traits as predictors of cooperation in a repeated prisoner's dilemma'. *Journal of Behavioral and Experimental Economics.* 64: 50-60

Anderson, S.P., Goeree, J.K., and Holt, C.A. (1998). 'A theoretical analysis of altruism and decision error in public goods games. *Journal of Public Economics.* 70: 297-323

Andreoni, J. (1988). 'Why free ride? Strategies and learning in public goods experiments'. J*ournal of Public Economics.* 37: 291–304

Andreoni, J. and Croson, R. (2001). 'Partners versus Strangers: Random Rematching in Public Goods Experiments'. *Handbook of Experimental Economics Results*. 1(6): 776-783

Andreoni, J., and Miller, J. (1993). 'Rational Cooperation in the Finitely Repeated Prisoner's Dilemma'. *The Economic Journal.* 103 (418): 570-585

Andreoni. J. (1988). 'Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence'. *The Journal of Political Economy.* 97(6): 1447-1458

Asongu, S., and Gupta, R. (2015). 'Trust and Quality of Growth: A Note'. Available at: http://www.afridev.org/RePEc/agd/agd-wpaper/Trust-and-Quality-of-Growth.-A-Note.pdf

Audretsch, D.B., Seitz, N., and Rouch, K.M. (2018). 'Tolerance and innovation: the role of institutional and social trust'. *Eurasian Business Review.* 8: 71-92

Australian National Audit Office (2009). 'Innovation in the Public Sector: Enabling Better Performance, Driving New Directions'. Available at: https://nla.gov.au/nla.obj-494725969/view

Bandura, A. (1986). *Social foundations of thought and action: Social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall

Baregheh, A., Rowley, J., and Sambrook, S. (2009). 'Towards a multidisciplinary definition of innovation'. *Management Decision.* 47 (8): 1323-1329

Battigalli, P., and Dufwenberg, M. (2007). 'Guilt in Games'. *American Economic Review.*97(2): 170-176

Becker, G.S. (1974). 'A Theory of Social Interactions'. *The Journal of Political Economy.* 82(6): 1063-1093

Becker, S.W., and Whisler, T.L. (1967). 'The Innovative Organization: A Selective View of Current Theory and Research'. *The Journal of Business.* 40 (4): 462-469

Ben-Ner, A., Kong, F., and Putterman, L. (2004). 'Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving'. *Journal of Economic Psychology.* 25(5): 581-589

Berg, J., Dickhaut, J., and McCabe, K. (1995). 'Trust, Reciprocity and Social History'. *Games and Economic Behaviour*. 10 (1): 122-142

Bialer, I. (1961). 'Conceptualization of success and failure in mentally retarded and normal children' *Journal of Personality*. 29: 303-320

Bidault, F., and Castello, A. (2010). 'Why too much trust is death to innovation'. *MIT Sloan management review*. Cambridge, Mass, 51(4): 33-38, 2010. Available at: https://www.researchgate.net/publication/290024720_Why_Too_Much_Trust_Is_Death_to_Innovation

Bjørnskov, C. (2005). 'The Determinants of Trust'. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=900183

Boone, C., De Brabander, B., and van_Witteloostuijn, A. (1999). 'The impact of personality on behavior in five Prisoner's Dilemma games.' *Journal of Economic Psychology.* 20: 343-377

Boone, C., Declerck, C., & Kiyonari, T. (2010). 'Inducing cooperative behavior among proselfs versus prosocials: The moderating role of incentives and trust'. *Journal of Conflict Resolution. 54*(5): 799–824.

Boone, C., De Brabander, B., and van Witteloostuijn, A. (1999). 'Locus of control and strategic behaviour in a prisoner's dilemma game'. *Personality and Individual Differences*. 27(4): 695-706

Bourdieu, P. (1986). 'The Forms of Capital'. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York: Greenwood Press

Brockman, P., Khurana, I.K. and Zhong, R.I. (2018). 'Societal trust and open innovation'. *Research Policy, Elsevier*. 47(10): 2048-2065

Bruttel, L., et al. (2023). 'Measuring strategic-uncertainty attitudes'. *Experimental Economics.* 26(3): 522-249

Budunchi, R. (2013). 'Trust, partner selection and innovation outcome in collaborative New Product Development'. *Production, Planning & Control.* 24(2,3): 145-157

Burnham, T., McCabe, K., and Smith, V. L. (2000). 'Friend-or-foe intentionality priming in an extensive form Trust Game'. *Journal of Economic Behavior and Organization*. 43: 57–73

Capraro, V., Jordan, J.J., and Rand, D, G. (2014). 'Heuristics guide the implementation of social preferences in one-shot Prisoner's Dilemma experiments'. *Scientific Reports.* (4). Available at: https://www.nature.com/articles/srep06790

Carpenter, J., and Huet-Vaughn, E. (2019). 'Real-effort tasks'. *Handbook of Research Methods and Applications in Experimental Economics*. (5)

Chami, R., and Fullenkamp, C. (1999). 'Trust and efficiency'. *Journal of Banking and Finance.* Elsevier, 26 (9): 1785-1809

Chaudhuri, A., Sopher, B. and Strand, P. (2002). 'Cooperation in social dilemmas, trust and reciprocity'. *Journal of Economic Psychology*. 23(2): 231-250

Ciccarone, G., Di Bartolomeo, G., and Papa, S. (2020). 'The rationale of in-group favoritism: An experimental test of three explanations'. *Games and Economic Behavior.* 124: 554-568

Clegg, C., et al. (2002). 'Implicating trust in the innovation process'. *Journal of Occupational and Organisational Psychology*. 75 (4): 409-422

Cohen, S. G., Ledford, G E., and Spreitzer, G M. (1996). 'A predictive model of self-managing work team effectiveness'. *Human Relations.* 49: 643-676

Coleman, J.S. (1990). *Foundation of social theory*. Cambridge, MA, and London, The Bellknap Press of Harvard University Press: 5-7

Confederation of British Industry (CBI)/QUINETIQ. (2008). *Excellence in Service Innovation*. CBI

Cosmides, L., and Tooby, J. (1989).'Evolutionary Psychology and the Generation of Culture, Part II'. *Ethology and Sociobiology.* 10: 51-97

Costello, T. & Prohaska, B. (2013). 'Innovation.' *IT professional*, 15 (3): 62-64

Covey, S.M.R. (2006). *The speed of trust: the one thing that changes everything.* Simon and Schuster

Cox, J.C. (2004). 'How to identify trust and reciprocity'. *Games and Economic Behavior.* 46(2): 260-281

Crosetto, P., and Filippin, A. (2016). 'A theoretical and experimental appraisal of four risk elicitation methods'. *Experimental Economics*. 19: 613-641

Croson, Rachel T.A. (1996). 'Partners and strangers revisited'. *Economics Letters.* 53: 25–32

Dakhli, M., and De Clercq, D. (2004). 'Human capital, social capital, and innovation: a multi-country study'. *Entrepreneurship & Regional Development.* 16 (2): 107-128

Damanpour, F. and Schneider, M. (2009). ‚Characteristics of Innovation and Innovation Adoption in Public Organisations: Assessing the Role of Managers'. *Journal of Public Resources and Theory*. 19: 495-522. Available at: https://academic.oup.com/jpart/article/19/3/495/940076

De Clercq, D., Thongpapanl, N., and Dimov, D. (2009). 'When good conflict gets better and bad conflict becomes worse: the role of social capital in the conflict–innovation relationship'. *Journal of the Academy of Marketing Science.* 37: 283-397

De Cremer, D. (1999) 'Trust and fear of exploitation in a public goods dilemma'. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*. 18(2): 153–163

De Cremer, D., and VanVugt, M.. (1999).'Social identification effects in social dilemmas: A transformation of motives'. *European Journal of Social Psychology.* 29(7): 871-893

De Vires, H.A., Bekkers, V.J.J.M., and Tummers, L.G. (2014). 'Innovations in the Public Sector: A Systematic Review'. *Ottawa IRSPM Conference*, 9-11 April 2014

Delle Foglie, L., Papa, S., and Spagnolo,G. (2025). 'Rage Against the Machine or Humans?'. *CEIS Working Paper No 593.* 23(1). Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5147649

Di Bartolomeo;G., and Papa, S. (2016a). 'Trust and reciprocity: extensions and robustness of triadic design'. *Experimental Economics.* 19 (1):100-115

Di Bartolomeo, G., and Papa, S. (2016b). 'Some Determinants of Trust Formation and Pro-social Behaviours in Investment Games: An Experimental Study'. *Studies in Microeconomics.* 4(1): 13-26

Dindaroglu, B. (2023). 'Interpersonal trust, invention, and innovation across European regions'. *Growth and Change. A Journal of Urban and Regional Policy.* 54 (3): 794-817

Drucker, P.F. (1969). *The age of discontinuity: guidelines to our changing society.* New York: Harper & Row.

Dunn, J.R., and Schweitzer, M.E. (2005). 'Feeling and Believing: The Influence of Emotion on Trust'. *Journal of Personality and Social Psychology.* 88(5): 736-748

Dunning, D., Fetchenhauer, D., and Schlösser, T. M. (2012). ,Trust as a social and emotional act: Noneconomic considerations in trust behavior'. *Journal of Economic Psychology.* 33 (3): 686–694. Available at: https://doi.org/10.1016/j.joep.2011.09.005

Dutta, S. et al. (2022). 'Global Innovation Index 2022: What is the future of innovation-driven growth?'. *World Intellectual Property Organization (WIPO)*

El Ghoul et al. (2023). 'Social trust and firm innovation'. *Journal of International Financial Markets, Institutions & Money*

Ellsberg, D. (1961). 'Risk, ambiguity and the Savage axioms'. *Quarterly Journal of Economics*. 75(4): 643–669.

Erikson, E.H. (1950). *Childhood and Society.* New York, NY, Norton & Company, Inc.

European Social Survey. (2002-2014). Available at: https://ess.sikt.no/en/?tab=overview

Evans, A.M. and Krueger, J.I. (2009). 'The psychology (and economics) of trust'. *Social and Personality Psychology Compass*. 3(6): 1003–1017

Evers, A., Ewert, B. and Brandsen, T., Eds. (2014). 'Social Innovations for Social Cohesion: 77 Cases from 20 European Cities 2010-14'. *European Commission*, Nijmegen.

Fischer, I. (2009). 'Friends or Foe: Subjective Expected Relative Similarity as a Determinant of Cooperation'. *Journal of Experimental Psychology: General.* 138(3): 341-350

Fisher, I. and Savranevski; L. (2023). 'The effect of similarity perceptions on human cooperation and confrontation'. *Scientific Reports.* 13. Available at: https://www.nature.com/articles/s41598-023-46609-8

Foxon, T.J. (2014). 'Technological lock-in and the role of innovation,' in: Aktinson, G., et al. (2007). *Handbook of Sustainable Development*, 20: 304-316. Edward Elgar Publishing.

Frisch, D., and Baron, J., (1988). 'Ambiguity and Rationality'. *Journal of Behavioral Decision Making.* 1(3): 149-157

Fukuyama, F. (1995). *Trust. The Social Virtues and the Creation of Prosperity*, New York, The Free Press. Chapters 3,4,7

Gintis, H., et al. (2002). 'Explaining altruistic behavior in humans'. *Evolution and Human Behavior.* 24:153-172

Grootaert, C., and Van Bastelar, T. (2002). *Understanding and Measuring Social Capital : A Multidisciplinary Tool for Practitioners*. World Bank Publications - Books, The World Bank Group, number 14098, December. Available at: https://ideas.repec.org/b/wbk/wbpubs/14098.html

Guilfoos, T., and Kurtz, K.J. (2017). 'Evaluating the role of personality trait information in social dilemmas'. *Journal of Behavioral and Experimental Economics.* 68:119-129

Gulati, R. (1998). 'Alliances and networks. *Strategic Management Journal.* 19(4): 293-317

Hasan, I., and Tucci, C. (2010). 'The innovation-economic growth nexus: Global evidence'. *Research Policy.* 39(10): 1264-1276

Hayashi, N., et al. (1999). 'Reciprocity, Trust and the Sense of Control'. *Rationality and Society.* 11(1): 27-46

Heidenreich, S., and Kraemer, T. (2015). 'Passive innovation resistance: The curse of innovation? Investigating consequences for innovative consumer behavior'. *Journal of Economic Psychology.* 51: 134–151

Henkel, L. (2024). 'Experimental evidence on the relationship between perceived ambiguity and likelihood insensitivity'. *Games and Economic Behavior.* 145: 312-338

Herbig, P.A., and McCarty, C. (1993). 'The Innovation Matrix'. *Journal of Global Marketing.* 6(4): 69-90

Heymann, J., and Ariely, D. (2006). 'Effort for payment: A tale of two markets'. *Psychological Science*. 15: 787–793

Hunt, S. (2012). 'Trust, Personal Moral Codes, and the Resource-Advantage Theory of Competition: Explaining Productivity, Economic Growth, and Wealth Creation'. *Contemporary Economics.* 6(2): 4-19

Jehn, K.A., and Mannix, E.A. (2001). 'The Dynamic Nature of Conflict: A Longitudinal Study of Intragroup Conflict and Group Performance'. *The Academy of Management Journal.* 44(2): 238-251

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., Knetsch, J.L., and Thaler, R. (1986). 'Fairness as a Constraint on Profit Seeking: Entitlements in the Market'. *The American Economic Review.* 76(2): 728-741

Kiyonari,T., Tanida, S., and Yamagishi,T. (2000). 'Social exchange and reciprocity: Confusion or a heuristic?'. *Evolution and Human Behavior.* 21(6): 411-427

Knack, S. (2002). 'Social Capital and the Quality of Government: Evidence from the States'. *American Journal of Political Science.* 46(4): 772-785

Knack, S., Keefer, P. (1997). 'Does Social Capital Have an Economic Payoff? A Cross-Country Investigation'. *Quarterly Journal of Economics*. 112(4): 1251-1288

Kosfeld, M., et al. (2005). 'Oxytocin increases trust in humans'. *Nature*. 435: 673-676

Krueger, J. I., Massey, A. L., and DiDonato, T. E. (2008). 'A matter of trust: From social preferences to the strategic adherence to social norms'. *Negotiation and Conflict Management Research*. 1: 31–52

Kuczmarksi, T.D. (2003). 'What is innovation? And why aren't companies doing more of it?'. *Journal of Consumer Marketing*. 20(6): 536-541

Langfred, C. W. (2004). 'Too Much of a Good Thing? Negative Effects of High Trust and Individual Autonomy in Self-Managing Teams'. *Academy of Management Journal.* 47(3): 385–399

Lazány, K. (2017). 'Innovation: the role of trust'. *Serbian Journal of Management.* 12 (2): 329-341

Leon, V., Etesami, S.R., and Nagi, R. (2022). Diffusion of Innovation under Limited-Trust Equilibrium. *2022 IEEE 61st Conference on Decision and Control (CDC).* December 6-9, 2022. Cancún, Mexico

LePine, J.A., and Van Dyne, L. (2001). 'Voice and Cooperative Behavior as Contrasting Forms of Contextual Performance: Evidence of Differential Relationships With Big Five Personality Characteristics and Cognitive Ability'. *Journal of Applied Psychology.* 86(2): 326-336

Li. C., Turmunkh, U., and Wakker, P.P. (2019). 'Trust as a decision under ambiguity'. *Experimental Economics.* 22(1): 51-75

Lyu, X., Ma, J., and Zhang, X. (2022). 'Social trust and corporate innovation: An informal institution perspective'. *North American Journal of Economics and Finance.* 64

Maccurtain, S., et al. (2008). 'Top Team Trust, Knowledge Sharing and Innovation'. Available at: https://core.ac.uk/download/pdf/11308821.pdf

Malhotra, D. (2004). 'Trust and reciprocity decisions: The differing perspectives of trustors and trusted parties'. *Organizational Behavior and Human Decision Processes*. 94: 61–73

Maradana, R.P., et al. (2017). 'Does innovation promote economic growth? Evidence from European countries'. *Journal of Innovation and Entrepreneurship.* 6(1)

March, J.G. (1991). 'Exploration and Exploitation in Organizational Learning.'. *Organization Science*, 2(1): 71-87

Markova, I., Linell, P. and Gillespie, A. ((2008). 'Trust and Distrust in society'. In Markova.I, Linell.P. e Gillespie, A. (eds). *Trust and Distrust: Sociocultural Perspectives.* Charlotte, N.C.: Information Age Publishing, pp.3-27

McCabe, K.A., and Smith, V.L. (2000). 'A comparison of naive and sophisticated subject behavior with game theoretic predictions'. *Proceedings of the National Academy of Sciences*. 97:3777–3781

McCabe, K.A., et al. (2001). 'A functional imaging study of cooperation in two-person reciprocal exchange'. *Proc Natl Acad Sci USA.* 98: 11832-11835

Meng, Y. et al. (2019). 'Trust and Corporate Innovation'. Available at SSRN: https://ssrn.com/abstract=3474170 or http://dx.doi.org/10.2139/ssrn.3474170

Mittone, L., and Papi, M. (2017). 'Does inducing choice procedures make individuals better off? An experimental study'. *Theory and Decision.* 83(3): 37-59

Mittone, L., Morreale, A., and Vu, T.-T -T. (2022). 'What drives innovative behaviour: An experimental analysis on risk attitudes, creativity and performance'. *Journal of Behavioural and Experimental Economics.* 98, p. 101868

Molina-Morales, F.X., and Martínez-Fernández, M.T. (2010). 'Social Networks: Effects of Social Capital on Firm Innovation'. *Journal of Small Business Management.* 48(2): 258-279

Molina-Morales, F.X., Martínez Fernández, M.T., and Torlò, V.J. (2011). 'The Dark Side of Trust: The Benefits, Costs and Optimal Levels of Trust for Innovation Performance'. *Long Range Planning.* 2: 118-133. Available at: https://www.sciencedirect.com/science/article/pii/S0024630111000033

Morris, S.D., and Klesner, J.L. (2010). 'Corruption and Trust: Theoretical Considerations and Evidence From Mexico'. *Comparative Political Studies*. *43*(10): 1258–1285

Morrison, K. and Pots, J. (2008). 'Toward Behavioural Innovation Economics – Heuristics and Biases in Choice under Novelty'. School of Economics Discussion Paper No. 379. Available at: https://ideas.repec.org/p/qld/uq2004/379.html

Mulgary, G., and Albury, D. (2003) Innovation in the Public Sector. Strategy Unit, Cabinet Office, London

Murtin, F. et al. (2018). 'Trust and its determinants: Evidence from the Trustlab experiment'. *OECD Statistics Working Papers*, 2018/02, OECD Publishing, Paris. Available at: http://dx.doi.org/10.1787/869ef2ec-en

NESTA (2012). 'Plan 1: The Case for Innovation Led Growth'. *NESTA*, London, 17

Neto J.C., et al. (2022). 'Confirmation Bias in Innovation Decision-Making'. *Revista Cientifica Multidisciplinar.* 3(7). Available at: https://dspace.uevora.pt/rdpc/handle/10174/32442

Neto, J.A.F., Felipe, J.A., and Bento Caleiro, A. (2019). 'Creativity and innovation: A contribution of behavioural economics.' *Psychology Compass*, 3(6): 1003–1017

OECD (2001). *The Wellbeing of Nations: The Role of Human and Social Capital, Education and Skills*. OECD Centre for Educational Research and Innovation, Paris, France

OECD (2017). 'OECD Guidelines on Measuring Trust'. *OECD Publishing*. Paris. Available at: https://www.oecd.org/governance/oecd-guidelines-on-measuring-trust-9789264278219-en.htm

OECD. (2005). 'The Measurement of Scientific and Technological Activities: Guidelines for Collecting and Interpreting Innovation Data: Oslo Manual, Third Edition'. Prepared by the Working Party of National Experts on Scientific and Technology Indicators, OECD, Paris, para. 146

Olson, J.A., et al. (2021). 'Naming unrelated words predicts creativity'. *Proceedings of the National Academy of sciences.* 118 (25)

Ortiz-Ospina, E., Roser, M., and Arriagada, P. (2016). 'Trust'. Published online at OurWorldInData.org. Available at: https://ourworldindata.org/trust

Palfrey, T.R., and Prisbey, J.E. (1996). 'Altrusim, reputation and noise in linear public goods experiments'. *Journal of Public Economics.* 61(3): 409-427

Pennington, R. Socher, C. D. Manning, 'GloVe: Global vectors for word representation'. *Empirical Methods in Natural Language Processing (EMNLP)* (Association for Computational Linguistics, Stroudsburg, PA 2014): 1532–1543

Poppo, L., Zhou, K.L. e Li, J.J (2016). 'When can you trust "trust"? Calculative trust, relational trust, and supplier performance.' *Strategic Management Journal*. 37(4): 724-741

Posen, H.E., and Levinthal, D.A. (2012). 'Chasing a Moving Target: Exploitation and Exploration in Dynamic Environments'. *Management Science.*58 (3): 587-661

Pots, J. (2010). 'Can behavioural biases in choice under novelty explain innovation failures?'. *Prometheus.* 28(2): 133-148

Powell, W.W., Koput, K.W., and Smith-Doerr, L. (1996). 'Interorganisational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology'. *Administrative Science Quarterly.* 41(1): 116-145

Pradhan, R.P., et al. (2016). 'Innovation, financial development and economic growth in Eurozone countries'. *Applied Economic Letters.* 23(16): 1-4

Putnam, R. (1993). *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton: Princeton University Press.

Rabin, M. (1993). 'Incorporating Fairness into Game Theory and Economics'. *The American Economic Review.* 83(5): 1281-1302

Rampersad, G., Genevieve Quester, P., and Troshani, I. (2010). 'Examining Network Factors: Commitment, Trust, Coordination and Harmony'. *Journal of Business & Industrial Marketing.* 25 (7): 487-500.

Rilling, J.K., et al. (2002). 'A Neural Basis for Social Cooperation'. *Neuron*. 35(2): 395-405

Romer, P.M. (1986). 'Increasing Returns and Long-Run Growth'. *The Journal of Political Economy.* 94(5): 1002-1037

Ross, L., Greene, D., and House, P. (1977). 'The false consensus effect: An egocentric bias in social perception and attribution processes'. *Journal of Experimental Social Psychology.* 13(3): 279-301

Rothstein, B., and Uslaner, E. (2005). 'All for One: Equality, Corruption, and Social Trust'. *World Politics.* 58: 41-72

Rousseau, D. M., et al. (1998). 'Not so different after all: A cross-discipline view of trust'. *Academy of Management Review*, 23: 393–404

Sankowska, A. (2013). 'Relationships between Organisational Trust, Knowledge Transfer, Knowledge Creation and Firm's Innovativeness'. *The Learning Organization.* 20(1): 85-100

Sapienza, P., Toldra-Simats, A., Zingales, L. (2013). 'Understanding Trust'. *The Economic Journal.* 123 (573): 1313-1332

Sarooghi, H., Libaers, D., and Burkemper, A. (2015). 'Examining the relationship between creativity and innovation: A meta-analysis of organizational, cultural, and environmental factors'. *Journal of Business Venturing.* 30(5): 714-731

Schneider, M., and Schields, T. (2022). 'Motives for Cooperation in the One-Shot Prisoner's Dilemma'. *Journal of Behavioral Finance*. 23(4): 438-456

Schroeder, K.B., Nettle, D., and Mc Elreath, R. (2015). 'Interactions between personality and institutions in cooperative behaviour in humans'. *Philosophical Transactions of the Royal Society B: Biological Sciences.* Available at: https://www.researchgate.net/publication/283453294_Interactions_between_personality_and_institutions_in_cooperative_behaviour_in_humans

Schumpeter, J.A. (1934). *Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle.* Cambridge, MA: Harvard University Press

Schumpeter, J.A. (1942). *Capitalism, Socialism and Democracy.* New York, Harper and Brothers

Selten, R. (1967). 'Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments'. *Beiträge zur experimentellen Wirtschaftsforschung*: 136-168

Shazi, R., Gillespie, N. and Steen, J. (2014). 'Trust as a predictor of innovation network ties in project teams'. *International Journal of Project Management.* 33: 81-91

Solow, R.M. (1956). 'A contribution to the Theory of Economic Growth'. *The Quarterly Journal of Economics.* 70 (1): 65-94

Steinbruch, F.K., da Silva Nascimento, L., Callegaro de Menezes, D. (2021). 'The role of trust in innovation ecosystems'. *Journal of Business & Industrial Marketing.* Available at: https://www.researchgate.net/publication/350700224_The_role_of_trust_in_innovation_ecosystems

Sternberg, R. J., and Lubart, T. I. (1999). 'The concept of creativity: Prospects and paradigms. In R. J. Sternberg (Ed.)'. *Handbook of creativity*: 3–15. Cambridge University Press

Suneja, V., and Das, D. (2024). 'Impact of Social Preferences, Trust and Behavioural Norms on Cooperation: Experimental Evidence from Prisoner's Dilemma Game'. *Studies in Microeconomics.* 12(3): 326-344

Taylor, S.P. (2017). 'What is innovation? A study of the definitions, academic models and applicability of innovation to an example of social housing in England'. *Open Journal of Social Sciences.* 5(11): 128-146

Teramoto, S. and Jurčys, P. (2017). 'Intermediaries and mutual trust: The role of social capital in facilitating innovation and creativity.' *New Technology, Big Data and the Law*: 129-150. Available at: https://link.springer.com/chapter/10.1007%2F978-981-10-5038-1_6

Thompson, M. (2018). 'Social capital, innovation and economic growth'. *Journal of Behavioral and Experimental Economics.* 73: 46-52

Tsai, W., and Ghoshal, S. (1998). 'Social Capital and Value Creation: The Role of Intrafirm Networks'. *The Academy of Management Journal.* 41(4): 464-476

Uzzi, B. (1996). 'The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect'. *American Sociological Review.* 61: 674-698

Van Lange, P.A.M., and Kuhlman, D.M. (1994). 'Social Value Orientations and Impressions of Partner's Honesty and Intelligence: A Test of the Might Versus Morality Effect'. *Journal of Personality and Social Psychology.* 67(1): 126-141

Watson Wyatt (2002). 'Human Capital Index: Human Capital As a Lead Indicator of Shareholder Value'. Available at: http://www.blindspot.ca/PDFs/HumanCapitalIndex.pdf

Yamagishi, T., and Kiyonari, T. (2000). 'The group as the container of generalized reciprocity.' *Social Psychology Quarterly.* 63(2): 116-132

Yusuf, S. (2009). 'From creativity to innovation'. *Technology in Society.* 31: 1-8

Zak, P.J. (2006). 'The Neuroeconomics of Trust' (2006). *Hendricks Symposium--Department of Political Science*. 9. Available at: https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1008&context=politicalscienceh endricks

Zak, P.J., and Knack, S. (2001). 'Trust and Growth'. *The Economic Journal.* 111 (470): 295-321. Available at: https://doi.org/10.1111/1468-0297.00609