Politecnico di Torino

Master's Degree Course in Engineering and Management

Academic Year 2022 / 23

Graduation session July 2023

How biases influence decision makers:

a focus on the behavioural variability in the Newsvendor Problem

Advisor:

Prof. Marco Cantamessa

Co-Advisors:

Prof. Giulio Zotteri

PhD Stud. Samuele Colombo

Candidate:

Francesca Tornabene

Abstract

During the 20th century, in a framework of growing opposition between Neoclassical and Behavioural Economics, it has been scientifically demonstrated that individuals cannot be considered perfectly rational and therefore capable of maximizing their utility when making choices, as commonly argued until then.

Rather, since the 1950s, various studies, in the context of Behavioural Economics, started showing that humans have inherent cognitive limitations that lead them to make systematically sub-optimal choices.

Since then, research has gradually uncovered multiple factors that influence decision makers in their choices because of these limitations, but the field continues still to be in the process of being explored and needs for additional contributions. Indeed, some aspects of the topic seem to be currently uncovered. The present study thus aims at investigating specifically one of these elements on which a lack of contributions is nowadays evident: the behavioural variability of decision makers.

To properly conduct this analysis, the Newsvendor Problem has been selected as the suitable tool, because of the many facets it has been proved to be able to investigate. As a matter of fact, the Newsvendor Problem reproduces a realistic decision making context, which can be easily linked to many daily activities actually carried out by individuals. Moreover, thanks to the versatility of the model, the behavioural variability was investigated in this work not only in the quantity purchased by the decision makers (i.e. the newsvendors), but also in their demand forecasting, costs incurred and deviation from optimal values (both in terms of costs and purchase quantity).

Through the use of the Analysis of Variance (ANOVA), a data analysis was thus conducted to study how behavioural variability in the above mentioned elements is impacted by different factors, such as the experimental characteristics and the cognitive abilities.

In particular, the results of the analysis showed how different combinations of these factors lead to diverse outcomes in the answering behaviour of the newsvendors, hence producing the investigated variability.

Furthermore, the present paper also aims at contributing to the study of behavioural variability in decision making through the proposal of an experiment, which has been designed to investigate the topic from an innovative biometric perspective. In this regard, even though for the purpose of the

present work the experiment has been conducted only in its pilot form, it is hoped that others will contribute to the research by going on in the data collection.

Summary

Index of Tables

Index of Figures

Introduction

The present document focuses on the behavioural variability characterizing individuals during decision making processes, with particular attention to the Newsvendor Problem. In detail, during the dissertation, the topic is analysed trough the results of a literature review and an analytical study, with the addition of an experimental proposal from a biometric perspective.

Problem formulation

Decision makers have not always been deemed narrowly rational individuals, who are influenced by a variety of circumstances in making their choices. Indeed, for a long time they were considered as perfectly rational and capable of always maximizing their utility. However, since the evidence of their limited rationality has been confirmed, numerous studies started being devoted to the issue, with the aim of understanding as many as possible facets of the topic.

Nevertheless, despite the well-established literature base, some elements of the question still need to be integrated. In detail, the conducted literature review revealed a lack of studies on the behavioural variability of decision makers.

The high variance characterizing decision makers' choices, which makes them systematically deviate from optimal practises, has been therefore deemed essential to study in this paper. However, the reasons for this choice reside beyond the evident shortage of contributions. Indeed, deepening the topic has been considered of fundamental relevance firstly because of its value in terms of practical applicability. Understanding the factors that make decision makers deviate from best practices when making choices is in fact the first step to be taken in order to improve corporate managerial performance, and reduce economic losses.

Specifically, some questions emerged from the conducted literature review:

1. *What is the relevance of behavioural variability in decision making?*
2. *What impact do human and cognitive characteristics have on behavioural variability?*
3. *What is the complexity of the different types of decisions?*
4. *Why is the Newsvendor Problem used to investigate behavioural variability in decision making?*
5. *What is already known about behavioural variability in decision making?*

Objective of the research

The present work aims to investigate possible causes of the behavioural variability that characterizes individuals when making choices, with a specific focus on understanding whether decision conditions, individual cognitive abilities and previous period conditions determine the variance of the behaviour of decision makers. In this regard, the final scope of the study is hence to define specific conditions that can help individuals to not deviate from optimal behaviours, in order to describe some potential settings that could help improving managerial decisions and reducing economic losses.

To this purpose, the conducted literature review revealed the Newsvendor Problem to be the most suitable tool to represent many of the daily structured decision-making processes undertaken by managers.

Moreover, given the current exploratory nature of the topic, this work also aims at providing an innovative experimental setting that could support the investigation of the phenomenon from a new challenging biometric perspective, which could be valuable to understand variability from the highlighting of the human mechanisms behind the decision making process.

Methodology

The present study was inspired by a review of the literature on decision making, which revealed the presence of many biases affecting decision makers, currently being deepened and contributed.

The exploratory nature of the topic thus gave rise to this work, which, through a mapping of decision making determining characteristics, found in the Newsvendor Problem the most suitable instrument of analysis.

From this consideration, the idea of assessing the unexplored behavioural variability characterising decision makers in the Newsvendor Problem context was then formulated, proceeding with an analytical analysis on a previously collected database, and through the proposal of an original biometric experimental setting.

During each phase of the study process, based on the available literary results, an attempt was made to extract a neutral definition for each topic, with the aim of conducting an as much as possible unbiased and objective analysis.

In detail, Pico, Scopus and Google Scholar were used to carry out the study properly. Specifically, the search engines were employed to realize the research on decision making and the Newsvendor Problem, thus to base the experimental and analytical subsequent phases. In this regard, the scientific articles to be considered were always carefully reviewed and ascertained in validity, specifically based on their diffusion in the scientific context.

Instead, for what properly concerns the analytical phase, many ANOVAs, T-tests, Mann-Whitney tests, and Kruskal-Wallis tests, with annexed graphical analysis, were realized through the IBM SPSS software.

Structure of the document

The document consists of six chapters, with the addition of the present introduction and the conclusions drawn.

The first chapter focuses on the conducted literature review on decision making and annexed biases affecting individuals during the process.

The second chapter is instead specifically dedicated to discussing studies on the Newsvendor Problem.

Then, the third, fourth, and fifth sections of the paper concern the analytical study conducted. In detail, the third chapter concerns the research questions and hypothesis that guided the analysis. The fourth part focuses on the data collection (carried out by researchers in 2019) on which the analysis was based. Still, the fifth chapter explains the data analysis realized specifically for this thesis.

The sixth chapter describes instead the biometric Newsvendor Problem experiment realized in its pilot form for this dissertation

Finally, in the conclusions, an overview of what has been uncovered in the paper is reported.

1. Decision making

"Whatever the Manager does he does it through setting objectives and making decisions aimed at their attainment", famously said the economist Drucker [1].

Decision making relevance in the management field is unanimously undoubted. It is one of the most critical tasks managers undertake at all levels and types of businesses. It is the act that most distinctly distinguishes management from other professions in society as a whole [2]. The Nobel Prize for Economics Simon considers it synonymous with the entire management action [3]. "Decisions are the core transactions of organizations. Successful companies "out decide" their competitors in at least three ways: they make better decisions; they make decisions faster; and they implement decisions more", highlights McLaughlin [4].

However, defining the concept independently of any perspective is first fundamental to start a dissertation about decision making relevance. To this end, for this thesis, many definitions were reviewed before defining the "zero-level" one, each focusing on different aspects of the topic. Thus, following some of the principal ones considered are reported, in addition to those already mentioned.

Starting from Stacey, with a strategic viewpoint, decision making is the process by which management determines and selects alternative courses of action appropriate to the circumstance's demands [5].

From a psychological perspective, Ajzen describes decision making as a subjective action based on the formulation of beliefs, with these being the subjective likelihood that an object possess a particular attribute [6] [7].

Orasanu and Connolly describe it as a sequence of conscious cognitive processes that take into account the characteristics of the surroundings [8].

Still, according to Narayan and Corcoran-Perry, making decisions entails the interplay of an issue that requires to be solved and someone who wants to solve it in a certain setting [9].

Finally, taking into account a fully-rational decision making process, it is considered that decisions are made after carefully considering several separate phases that are planned and that can be examined logically and impartially [10] [11].

In conclusion, after having analysed them, the many decision making definitions were summarized in the subsequent one, which aims at extrapolating the concept from any meaning and any peculiar focus: "The choice of a course of action among different alternatives". This definition represents how the present work conceives decision making.

1.1. Approach over time

Across the years, many different approaches have characterized the study of two intrinsically connected fields, namely economics and decision making, as summarized in the following lines, which attempt to provide an overview of the way in which the conceptions about the topic have changed over time.

Starting from the Normative Theory of decision making in economics, its foundations date back to when Samuelson in 1938 identified utility as the total of choices represented in a person's conduct [12]. In 1947 the author then constituted the basis of Neoclassical economics by delineating multiple hypotheses regarding individuals conduct when taking a decision. In particular, Samuelson stated that decision makers must be aware of all the possible choices, have clear ideas regarding the effects of their choices, and be guided by the principles of rationality, self-interest, and logic [13].

In that context, characterized by the observation of the decision maker's perfect rationality, Simon introduced the "Homo economicus" concept, representing the perfect economic man who "... is assumed to know the relevant aspects of his environment which, if not complete, is at least impressively clear and voluminous". However, the economist criticised applying this idea in economics theory, underlying that decision makers should be considered as bounded rational individuals who attempt to be satisfied more than maximize their utility [14].

In this regard, a few years later, Kahneman and Tversky revealed the importance of these elements, showing that humans systematically deviate from optimal prescribed behaviours when taking a decision. Also, through the years, the authors developed a "map of bounded rationality" by analysing the gaps, the so-called biases, that distinguish the rational-agents model prescriptions from reality [15].

However, despite the huge amount of studies regarding the field of human decision making, there is no lack of conflicting opinions on the subject. Among the most common criticisms, the randomness of the behavioural deviations has been addressed by many, who postulate that this is why rationality rules can be neglected [16]. In addition, the learning ability of individuals has been considered to

solve the irrationality issue, claiming that people can improve in performance from their mistakes. However, these observations have been replicated by highlighting the massive level of behavioural deviations, defined by Kahneman and Tversky: "too widespread to be ignored, too systematic to be dismissed as random error, and too fundamental to be accommodated by relaxing the normative model" [17].

In conclusion, it can be asserted that the main difference in the proposed approaches lies in the distinct way psychologists and economists face problems and drive models and conclusions from them. While economic models are generally formal and precise, behavioural formulations tend to leave more room for doubts and uncertainties, sometimes more consistent with the many facets of reality.

1.2. Decision biases

Decision biases have produced a vast interdisciplinary debate among economists and psychologists over the years.

The reason for the researchers' in-depth investigation of these phenomena lies in the role that these elements play in the understanding of decisions, which made it of great interest for the present study to contribute to the field through a focus on how decision biases influence the decision making process. In fact, biases caution analysts against blindly embracing intuitive judgements made by decision makers and highlight the importance of assisting decision makers in avoiding them [18].

However, to better understand these points, it is first necessary to define the "zero-level" concept of cognitive bias. For the present work, many interpretations were therefore reviewed to reach the most aseptic one, both in terms of cognitive and decision biases. The two terms appear, in fact, deeply connected and sometimes treated as interchangeable, but their relationship is hierarchical in terms of abstraction: decision making bias can be said to be the cognitive bias applied to decisions, as it is evident in the subsequent lines, where some of the most significant definitions are presented.

Haselton, Nettle, and Andrews defined cognitive biaseas "*the most dramatic demonstrations of the role of the mind in our apprehension of the world",* with the addition that *"where biases exist, individuals draw inferences or adopt beliefs where the evidence for doing so in a logically sound manner is either insufficient or absent*" [19].

According to Barberà-Mariné et al., the meaning of cognitive biases is instead that of "*useful measurements for detecting process improvement actions*" [20].

Finally, Arnott defined cognitive biases as "predictable deviations from rationality", highlighting that "they are inherent in human reasoning" and aggregating the two concepts of biases to the point of asserting that cognitive biases are often referred to as decisions [21].

On the other hand, regarding decision biases, the same Arnott, in another paper, defined decision biases as "deviations from rational decision making", specifying that they "are often positive manifestations of human information processing ability, as the action of biases is generally a source of decision effectiveness", but also underlying their role in causing "poor decision outcomes in important decisions" [22].

Still, in 2003 Kahneman wrote about the "systematic biases that separate the beliefs that people have and the choices they make from the optimal beliefs and choices assumed in rational-agent models" [15].

Concluding, more recently, in 2015, Montibeller and von Winterfeldt focused on the cognitive biases relevant for decision and risk analysis, "the faulty mental processes that lead judgments and decisions to violate commonly accepted normative principles", that "can distort the results of an analysis and are difficult to correct" [23].

Also just from these definitions it is clear what cognitive biases represent:" Thinking deviations from rationality". Therefore, by applying this concept to decision making, also given what was stated in the previous paragraphs, the researched neutral decision bias definition is obtained as "Choices deviations from rationality". These are the definitions to be considered for the analysis of this work.

The analysis of these biases, of how they are inherent to human beings, has pervaded many fields over the years, assuming different denominations depending on the area of investigation, being referred to as "behavioural finance" in finance [24], "judgement and decision making" [25] and "the psychology of decision making" [26] in management and industrial psychology.

1.2.1. Overview of the biases

The topic is so wide and facetted that a plethora of cognitive biases has emerged over the years as resulting from a literature review in different domains, such as psychology, economics and organizational decision making. The research field is so broad and constantly evolving that it is quite

impossible to realize a comprehensive list of all the biases so far identified, which today number in the hundreds, if not higher in some studies. Just to have an idea of the magnitude of the topic, despite it is beyond the purpose of this work to provide an exhaustive review of all the existing biases (i.e., those so far identified), in Figure 1, an overview of 188 biases and heuristics, in turn, divided into subcategories, is reported (adapted from [27]).

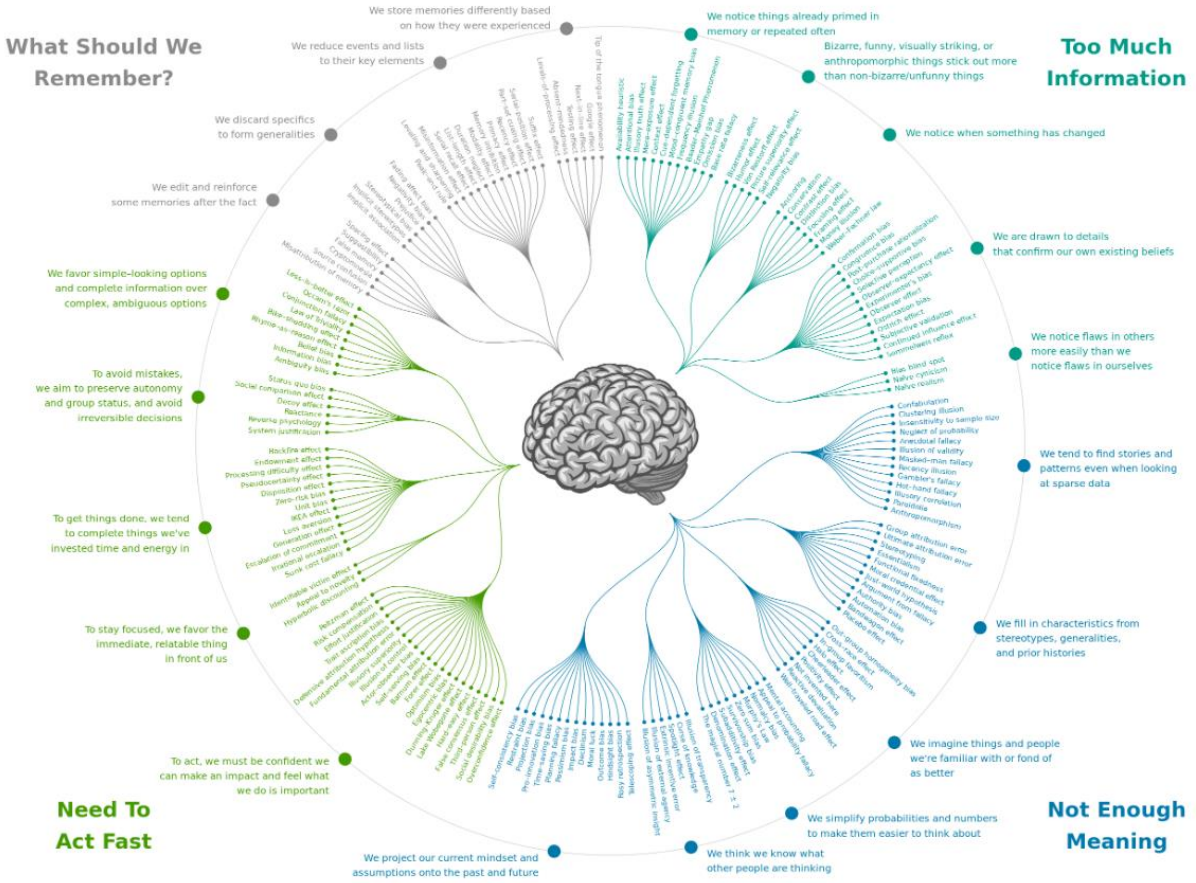

Figure 1: Cognitive bias codex

Among the already established cognitive biases, some have resulted particularly relevant in decision making processes. In this regard, an overview of those resulted as the most famous ones from the review is reported in Table 1.

Table 1: Most famous decision biases

DECISION BIAS	DESCRIPTION	REFERENCE
Anchoring and adjustment	Decision makers tend to anchor around a position and (insufficiently) adjust from this reference point	Tversky and Kahneman, 1974 [28]
Availability	Decision makers tend to utilize information based on how easily this is available	Tversky and Kahneman, 1974 [28]
Confirmation	Decision makers tend to search for confirmation and not for disconfirming information	Einhorn and Hogarth, 1986 [29]
Familiarity	Decision makers tend to consider more familiar facts as more probably to occur than less familiar ones	Fox and Levav, 2000 [30]
Law of small number	Decision makers tend to draw conclusions from inappropriately small samples	Tversky and Kahneman, 1971 [31]
Loss aversion	Decision makers tend to opt for avoiding losses than obtaining gains	Kahneman and Tversky, 1979 [32]
Overconfidence	Decision makers tend to underestimate the outcome of an uncertain event	Moore and Healy, 2008 [33]
Recency effect	Decision makers tend to attribute more weight to recent facts than earlier ones	Cushing and Ahlawat, 1996 [34]
Similarity	Decision makers tend to evaluate the probability of occurrence of an event based on the probabilities of similar ones	Kahneman and Tversky, 1973 [35]
Sunk costs fallacy	Decision makers tend to continue an investment once it has been made	Shaanan, 1994 [36]

1.2.2. Taxonomies of the biases

Considering the vastness of the identified biases in literature, a taxonomy of these is required to have an overview of how they can be grouped and how they differently influence decision making.

In the following lines, two of the "main" taxonomies are reported, representing what turned out as some of the most recognised ones from the literature review.

However, to ensure a critical reading of the presented taxonomies, it is first necessary to specify the characteristics that make a taxonomy valid from a scientific viewpoint. These particularly refer to the

requirements of internal consistency and of correct identification of abstraction levels. These qualifications can be reached by clustering biases, categorizing them into groups that internally share comparable characteristics and differentiate from others [37]. Moreover, it is also notable to mention that the currently existing classifications are mainly based on subjective criteria rather than rigid schemes.

Starting with the already mentioned and well-known Tversky and Kahneman, they offered a nowadays still authoritative taxonomy of decision biases, which, according to the authors, can be traced back to different heuristics. They explain how most important decisions are built on these heuristics, leading to systematic errors. In particular, the authors identified three main heuristics on which people rely while making choices, namely: Representativeness (i.e., the evaluation of the degree to which an object X is similar to an object Y), Availability (i.e., the ease with which instances a given topic can be thought) and Anchoring and Adjustment (i.e., making estimates starting from a value which is adjusted till the final one) [28]. Relying on these heuristics, the biases presented in Table 2 are produced.

Table 2: Tversky and Kahneman's Decision biases taxonomy

HEURISTIC	BIAS
Representativeness	*Insensitivity to prior probability of outcomes* *Insensitivity to sample size* *Misconceptions of Chance* *Insensitivity to predictive accuracy* *The illusion of validity* *Misconceptions of Regression*
Availability	*Biases due to the retrievability of instances* *Biases due to the effectiveness of a search set* *Biases of imaginability* *Illusory correlation*
Anchoring and Adjustment	*Insufficient adjustment* *Biases in the evaluation of conjunctive and disjunctive events* *Anchoring in the assessment of subjective probability distributions*

Remus and Kottemann distinguished instead decision biases basing on whether they concern data presentation to the decision maker or how information is processed. In this way, the authors aim at

underlying how information is first spontaneously distorted by the sensory input channels, which gather the necessary information for the decision maker to make their decision, and then again distorted in the phase of data processing as well. Also, the authors explain how these biases cannot be eliminated, even with training, because of the inherent neurophysiological limits of the human brain. These two types of biases are represented by many sub-biases, which in turn are generally further divided in other ones, respectively represented in the central and right column of Table 3 (please note that the distinction in "First", "Second" and "Third level of description" was realized for the purpose of this work and does not belong to the authors who realized the taxonomy) [38].

Table 3: Remus and Kottemann's Decision biases taxonomy

FIRST LEVEL OF DESCRIPTION	SECOND LEVEL OF DESCRIPTION	THIRD LEVEL OF DESCRIPTION
Biases associated with presenting data to a decision maker	*Irrelevant Information*	*/*
	Data Presentation Biases	*The type of information* *The format of the data display* *Logical data displays* *Order effects* *context*
	Selective Perception	*People filter data in ways that reflect their experience* *People's expectations can bias perceptions* *People seek Information consistent with their own views* *People downplay data which contradicts their views*
	Frequency	*Ease of recall* *Base rate error* *Frequency to imply strength of relationship* *Illusory correlation*

Biases in information processing	*Heuristics*	*Structuring problems based on experiences* *Rules of thumb* *Anchoring and adjustment* *Inconsistency in the use of the heuristic*
	Misunderstanding of the Statistical Properties of Data	*Mistaking random variation as a persisting change* *Inferring from small samples* *Gambler's fallacy* *Ignoring uncertainty*
	Limited Search Strategies	/
	Conservatism in decision making	/
	Inability to Extrapolate Growth Processes	/

1.3. Main drivers

To deeply understand the definition of decision making in all its possible nuances and the amplitude of the concept, it is necessary to present a deeper analysis of the main drivers of a generic decision, which embody the intrinsic determinants of choices on different levels.

An overview of the main elements that characterize decision making processes is hence reported in Table 4. All the identified elements were found to deeply influence each decision, by shaping it with many different facets.

The following partition into four main clusters and the subsequent division into subcategories were produced by aggregating literary results, but they do not belong to any specific author.

Table 4: Decision making main drivers

CLUSTER	SUBCATEGORY	DRIVER
Decision maker's personal traits	Non-ability factors	*Gender,* *Age,* *Personality,* *Interests,* *Culture,* *Risk preferences,* *System 1 and 2*
	Ability factors	*Intelligence,* *Experience*
	Incidental factors	*Emotions,* *Stress,* *Motivation,* *Sleep,* *Circadian timing*
Biological factors	/	*Weather,* *Time-of-day,* *Air pollution*
Task characteristics	Performance-independent factors	*Environment,* *Uncertainty*
	Performance-dependent factors	*Reward,* *Feedback*
Single decision characteristics	Decision maker-related factors	*Number of decision makers involved* *Need to generate solutions* *Information availability*
	Decision related factors	*Number of solutions* *Number of repetitions* *Link with other decisions* *Time horizon*

1.3.1. Decision maker's personal traits

The first group of drivers concerns decision makers' personal traits, subdivided into three main subcategories.

The first subcategory includes the so-called "Non-ability factors, " which disregard individuals' skills and are mainly influenced by social or environmental circumstances.

Gender can be mentioned as a relevant element in determining how individuals face decision making: studies have proven that, while women tend to be more influenced by their surroundings, spend

more time researching options and making decisions, men tend to be more forceful, powerful, objective, and practical [39] [40].

Contrasting opinions have emerged instead through the years about age influence on decision making: some authors think that there is a distinction in how people approach the process based on their age [41] [42], whereas others do not agree [43] [44]. However, the fact that more experience is gained with age should be considered when analysing decision makers of different ages [45].

In addition to the more "registry" data, individual emotional sphere has also been identified as impacting decision making. In this regard, the Big Five model of individual personality traits prescribes that one's personality is embedded in their openness to experience, conscientiousness, extroversion, agreeableness and neuroticism [46]. By fitting differently in various situations, all these elements deeply influence decision making. For instance, more open-to-new-experience people could be more appropriate for unstructured and innovative decision making tasks [47].

In accordance with what just stated about personality, decision making is also shaped by individual interests, synthesized by Holland theory in the following categories: realistic, investigative, artistic, social, enterprising, and conventional [48] [49]. These individual interests are considered as influencing decision making since they make individuals enjoy various situations differently. For instance, people with investigative interests might be more confident in decision making circumstances in which information is not already processed [47].

Moving on, to understand how distinct cultures face decision making processes, the "Cardinal Issue Perspective" (CIP) can be addressed [50]. The model identifies ten issues handled in almost every decision, that have been proven to be answered differently from different cultures:

- Does a decision need to be made or not?
- Who will make the decision, and how will they go about their work?
- Will the decider invest many or few resources into the decision making process?
- What options are available or could be created?
- What important possible outcomes are associated with each option?
- What is the likelihood of each outcome occurring?
- How good or bad would each outcome be for the decision maker (and/or other parties)?
- How should one manage trade-offs between options?
- What do other parties think of the decision?
- What can the decision maker do to ensure that the decision is implemented?

For instance, regarding the second issue, it has to be considered the difference between individualism-based and collectivism-based cultural approaches: in this context, in some studies Americans were less inclined than Russians to offer spontaneous advice, potentially reflecting a predisposition for individual decision making [51]. Generalising, it can be asserted that the type of culture (individualist or collectivist), the cultural norm strength (tight or loose) and the cognitive style (holistic or analytic) are some of the main drivers that differentiate the way in which these issues are faced from different cultures [52].

From another perspective, risk preferences have been identified as a discriminant in how individuals make decisions, describing the degree of risk that a person is ready to accept to achieve a goal. From this point of view, individuals can be divided into risk-averse and risk-seeking, with the former preferring less-risky options and the latter more inclined to make risky choices. Discussing risk preferences makes it inevitable to mention the Israeli psychologists Daniel Kahneman and Amos Tversky, authors of Prospect theory [32]. By building different experiments, they overcame Bernoulli's utility theory, noticing that "[...] the abrupt switch from risk aversion to risk seeking could not plausibly be explained by a utility function for wealth. Preferences appeared to be determined by attitudes to gains and losses, defined relative to a reference point" [15].

Finally, the last non-ability factor was identified in the concepts of System 1 and System 2, the metaphor passed from Stanovich and West, on which Kahneman's book "Thinking, Fast and Slow" focuses [53] [54]. The two Systems, which differ in domains by individuals, respectively represent: System 1 (or Thinking fast) is intuitive thinking. In contrast, System 2 (or Thinking slow) is a more effortful way of thinking, and it does not occur immediately. It has been proved that decision makers tend to rely on System 1, which generally leads to errors. However, people can improve their ability to switch from System 1 to System 2 to make better decisions. In conclusion, which of the two systems is engaged for the decision making process has a big impact on the choices that people make [55].

The second subcategory of decision maker's personal traits influencing decision making is related to individuals skills and experience, the "Ability factors".

In particular, about intelligence, the set of individual skills constituting one's cognitive ability, the PPIK theory from Ackerman distinguishes between "intelligence-as-process" and "intelligence-as-knowledge" (respectively, the first and the last letters of the acronymous), with the former including memory, reasoning, spatial rotation capabilities and perceptual speed, and the latter referring to the

competences that an individual acquires in a specific domain [56]. By applying these skills, individuals with higher intelligence tend to make more accurate and less biased choices [47].

Also, as just mentioned, experience itself plays a fundamental role in determining individual choices, as it is evident in the already given definition of "intelligence-as-knowledge": as one gains more experience, he/she also acquires more knowledge and so a better understanding of the decision making context and more confidence [56] [47].

Finally, the last subcategory of individual's personal traits influencing decision making is represented by the "Incidental factors" , those variable elements that characterise the individual when he/she has to make a decision.

Each of them is linked to a series of cognitive appraisals (the individual's subjective interpretations of environmental inputs) that direct their influence on decision making through psychological mechanisms. Depending on their experience and how they relate to the decision at hand, emotions can be distinguished in integral (integral to the decision making action) and incidental (arising in a particular situation, but not inherent in it). Both types of emotions have proved to influence decision making: it is sufficient to think about marketing, specifically designed as an inducer of emotions in consumers, to entice them to purchase [57].

Also, stress, "the non-specific response of the body to any demand for change", has proved to be an important determinant in decisions [58]. In particular, early studies demonstrate that stress impacts brain circuits involved in learning and reward processing [59] [60] and that it can also bias decisions toward habits [61] and modify the propensity to take risks [62] [63] [64]. Stress influence on the decision makers particularly depends on the type of stressor and on the stressor timing (stressor duration, stress-to-task latency and exposure across the lifespan) [65].

The role of incentives was investigated and it turned out that both internal (i.e., motivation) and external (i.e. explicit performance reward) incentives function in decision making has been widely studied [66]. For instance, in his review, Easterbrook compiled a large body of research demonstrating that humans only take into account a small number of cues in high drive phases (of motivation) [67].

Moreover, sleep loss was found to be relevant in the determination of one's choices too. It has negative effects on humans, as it can be seen both in simple psychomotor abilities (e.g. [68]) and in cognitive tasks [69]. Even if not all cognitive processes may be impacted by lack of sleep, decision making has been proved to be an exception. Sleep deprived individuals tend to favour risky choices (e.g. [70]), develop a desensitized attitude toward financial risk [71], tend to rely on straightforward

reinforcement-heuristic in a Bayesian decision environment [72] and decrease prosocial behaviours in short-term social interactions [73] [74]. Particularly, evidence suggests that complex decision making is deeply responsive to sleep deprivation [72] [75], which could be the cause of deficiencies in processes like attention (e.g., [76]) and updating (e.g., [77]) which are essential for making complex decisions.

Finally, the last incidental factor influencing decision making was found in the circadian timing (or rhythm), which represents the individual alterations in physiological and behavioural processes throughout the day. It has been proved that people engage in more "effortful thinking" when their circadian rhythm is on than when it is off (e.g. [78] [79]). Many studies have been conducted on this topic. For example, Bodehausen examined the potential impact of circadian match or mismatch on stereotype reactions, rationalizing that people experiencing a circadian mismatch should have reduced cognitive capacity, leading them to become less reflective and more dependent on heuristics and stereotypes. The study showed that participants were more prone to use stereotypes during circadian "off-times" than during circadian "on-times", when making decisions

1.3.2. Biological factors

The second category of identified decision making drivers refers to the biological factors characterizing the environment in which the decision has to be made.

Many studies have shown that weather can have a significant impact on a person's mood [80] and, in turn, on his/her decision making [81] and risk-taking [82]. In particular, lack of sunlight produces melatonin and cortisol disproportion, leading to weak energy, diminished optimism, and/or sadness. Also, sunshine and temperature impact hormone and neurotransmitter production, which can influence behaviour [83]. Specifically, lower temperatures can increase aggression [84] and higher temperatures are associated with a more pleasant mood and greater memory [85].

Modern studies have shown that individual decision making is also affected by the time of day and decision maker's chronotype, which represents whether he/she is a morning/evening person. The so-called "synchrony" effects [86] take place when the timing of optimal behavioural responses matches the phase of circadian arousal across individuals, i.e. when time of day and chronotype interact: morning-type individuals perform better at cognitive tasks in the morning and vice versa for evening-type individuals [87].

Moreover, with the nowadays increasing attention paid to pollution, studies have focused on the immediate adverse effect that this has an on individual's health and emotions. In turn, even if currently few contributions can be counted in the study of how pollution influences decision making, it has already emerged that air pollution directly influence risk and satisfaction conception [88]. Also, because of pollution, in post-decision evaluation, consumers tend to increase pessimistic feeling, decrease anticipated happy emotion, and provide lower review scores [89] [90] [91] [92].

1.3.3. Task characteristics

The third cluster of decisional process drivers is represented by the characteristics of the decision making task, grouped in two subsections.

This first subsection represents those factors that affect decision maker's behaviour regardless of his/her performance.

Among these elements, the environment was deeply relevant in the decision making process definition. In this regard, environments can be distinguished into stable and dynamic, depending on the level of uncertainty and changes that characterise them. For instance, it has been proved that dynamic environments hinder learning from experience, unlike the less realistic, stable ones, in which practicing and receiving feedback have demonstrated to be more effective [93].

Also, uncertainty, the absence of knowledge regarding the likelihood that an event will occur in the future [94] and the incapacity of a decision maker to totally eliminate it were identified as restricting decision making efficacy, necessitating the adoption of strategies that either serve to decrease or manage it. In detail, exogenous and endogenous uncertainty and environmental, industry, and company uncertainties are frequently distinguished in recent studies [95].

The second subgroup of decision making task features impacting decision making includes those elements that influence the decision maker consequently to his/her performance.

In this regard, what is probably most immediate to think about is a reward, all those external incentives that influence the behaviour of the decision maker based on his performance [66]. Results show that when people are given incentives, they spend more attention and time on the task and appear to be more "motivated". However, it is unclear whether this enhances performance and, more specifically, whether incentives lead people to focus on the proper dimensions of tasks. In general, experiment results imply that payoffs influence subjects' behaviour in the right direction, but people still do not behave as normative models require [96]. McCullers summarizes many studies on the

effects of incentives and concludes that they improve performance when it is dependent on making simple responses. He also asserts that the significance of incentives in activities that need flexible, open-ended, and creative solutions is significantly less evident [97].

On the other hand, giving feedback to the decision maker after his/her performance influences the decision making process from multiple perspectives too. For instance, increasing the frequency of feedback should improve performance since decision makers will be able to respond rapidly to changes in the environment and see the repercussions of their choices. However, since people often adapt their decision making processes to the information environment [98], more frequent feedbacks could worsen performance if they cause decision makers to focus on and process the latest data and overreact to random noise [99]. Also, studies have investigated how feedback strength [100], noise [101] and presence [102] influence performance and learning. In general, many recent studies have shown that the impacts of feedback [103].

1.3.4. Single decision characteristics

Finally, the fourth and last set of decision making drivers is represented by the attributes of a single, generic decision, identified in decision maker-related factors and purely decision-related elements.

As shown in Table 5, each of the following identified features was found to be fundamental in determining the type of decision. Two or more different types of decision making processes were in fact associated to each feature, based on its realisation.

Table 5: Decision attributes and deriving decisions

CLUSTER	ATTRIBUTE	TYPE OF DECISION
Decision maker-related factors	Number of decision makers involved	*Individual - Group*
	Need to generate solutions	*Structured - Unstructured*
	Information availability	*Uncertain - Ambiguous - Certain*
Decision related factors	Number of solutions	*Convergent - Divergent*
	Number of repetitions	*One-shot - Multistage*
	Link with other decisions	*Independent - Dependent*
	Time horizon	*Short term – Long term*

Starting with those characteristics of a decision that are in a certain way related to the decision maker his/herself, the first analysed attribute was the number of individuals making the decision. This feature gives rise to the distinction between “group” and “individual” decisions, with the former been

made jointly by more people and the latter by a single person. Currently contrasting studies dissert about which of these two could be addressed as the “more optimal” decision making process [104].

Moving on, based on the fact that the “best courses of action” already exist or have to be generated, decisions can be divided into “structured” and “unstructured”. This attribute also makes evident the distinction between the concepts of “problem solving” and “strategy”. In this regard, a common mistake regarding decision making is represented by its association with the concept of problem solving. This point can be clarified by looking at what Drucker observes: “[…] "problem-solving" assumes that we know what the problem is and have to find the one right answer to it the way engineering students, for instance, talk of "problems". But the first and most difficult job in decision making is always to find the right question. And there is rarely one "right answer" in a business situation; at best there is a range of alternatives, each with distinct risks of its own and none completely ''solving" the problem.” [1] As the author clearly describes, problem solving is only a specific type of decision making process, a structured and defined one, which is opposed to other different unstructured decision making forms.

The last decision making discriminant of the analysed category was found in the amount of information that the decision maker possesses, which differentiate decisions in “certain”, “uncertain” and “ambiguous”. While the distinction between certain and uncertain decisions could be easily thought, the one between uncertain and ambiguous might be more complex and it is not always made. The two attributes distinguish cases in which the elements that may be affected by uncertainty are clear and cases in which uncertainty occurs like the elements themselves [29].

Finally, looking at the purely decision-related, more “technical” features, the number of solutions admitted by a decision was one of the relevant identified factors, since it determines whether decision are “convergent” or “divergent”, allowing a single solution and more than one respectively [105] [106]. This distinction also leads to the differentiation among “optimization” and “heuristics”, two different types of strategies, attempting and not at optimizing [107].

Also, the number of repetitions characterized different processes: while in “one-shot” decisions, the decision maker only has one possibility to make his/her choice, differently, “multistage” decisions are made at different stages [108].

“Dependent” and “independent” decision definitions derive instead from whether they are constrained or not with previous choices. In particular, the concept of independent decisions is one

of the main pillars of the so-called Dynamic decision making, in which, as Edwards describes, an object is reached after a series of independent decisions [109].

To conclude, according to the extent of the time horizon in which the effects of the choices made by an individual fall, decision making is differentiated into “short-term” and” long-term”, with the former having an immediate and punctual impact on the surroundings and the latter a more diluted effect over time.

2. The Newsvendor Problem

For the present study, out of the many possible variants of decision making, one was found to be optimal for a comprehensive and focused exploration of how biases intervene in the decision making process, producing the investigated behavioural variability.

Table 6: Observed type of decision

CLUSTER	ATTRIBUTE	OBSERVED DECISION
Decision maker-related factors	Number of decision makers involved	*Individual*
	Need to generate solutions	*Structured*
	Information availability	*Uncertain*
Decision related factors	Number of solutions	*Convergent*
	Number of repetitions	*Multistage*
	Link with other decisions	*Independent*
	Time horizon	*Short term*

In detail, the attributes choice made (reported in Table 6) converged in the observation of one specific type of decision making example, a classic operations management model: the Newsvendor Problem.

Also referred to as the Newsboy Problem, the model entails a decision maker who sells a product during a short selling season in presence of stochastic demand, with the possibility to order inventory before the selling season only. In his/her choice, the decision maker has to calibrate the costs of ordering too little and too much, given that if the order quantity exceeds the actual demand, the manager must sell the leftover stock at a loss, and if it is less than the actual demand he/she forfeits some profit [93].

2.1. Definition of the model

Before delving further into the model, it is fundamental to identify it independently of any perspectives. This activity is conducted in the following lines by analysing different definitions belonging to various authors.

Following two of the most recognized researchers in the field, Schweitzer and Cachon, to describe the Newsvendor Problem from a purely mathematical perspective, it would be sufficient to identify it with a model in which a decision maker orders a quantity q before the selling period, at a unitary

cost c, in the presence of the stochastic demand D, characterized by a mean μ, a distribution function F, and a density function f. The newsvendor then sells each unit at a cost $p > c$ and salvages the remaining units for $s < c$, in case $q > D$, realising the profit

$\pi(D, q) = (p\text{-}c) \cdot \min(q, D) - (c - s) \cdot q$

with an expected profit of

$\mathbb{E}(\pi(q, D) = \big(1 - F(q)\big) \cdot \pi(q, q) + \int_0^q f(x) \cdot \pi(q, x) dx$

Moreover, being

$q_n = \arg\max \mathbb{E}\,[\pi\,(q, D)]$

the unique solution to

$F(q_n) = \frac{p-c}{p-s}$

the ratio $\frac{p-c}{p-s}$ is defined "critical fractile" and it is used to classify products in high profit and low profit, depending on whether

$\frac{1}{2} \leq \frac{p-c}{p-s}$

From a behavioural perspective, the Newsvendor Problem could be instead described as a model that produces recurrent divergences from the optimal behaviour, as explained by Bostian, Holth and Smith [110].

In this regard, to better clarify how the process works, its complexity from a behavioural point of view, and the dynamics that cooperate with the definition of the ordering decisions, an overview of this is reported in Figure 2 (adapted from [111]). In detail, the scheme, belonging to Yamini, offers a comprehensive view of how different cognitive biases, learning dynamics, bounded rationality, and the use of cognitive heuristics as thumb rules lead newsvendors to make wrong, biased decisions .

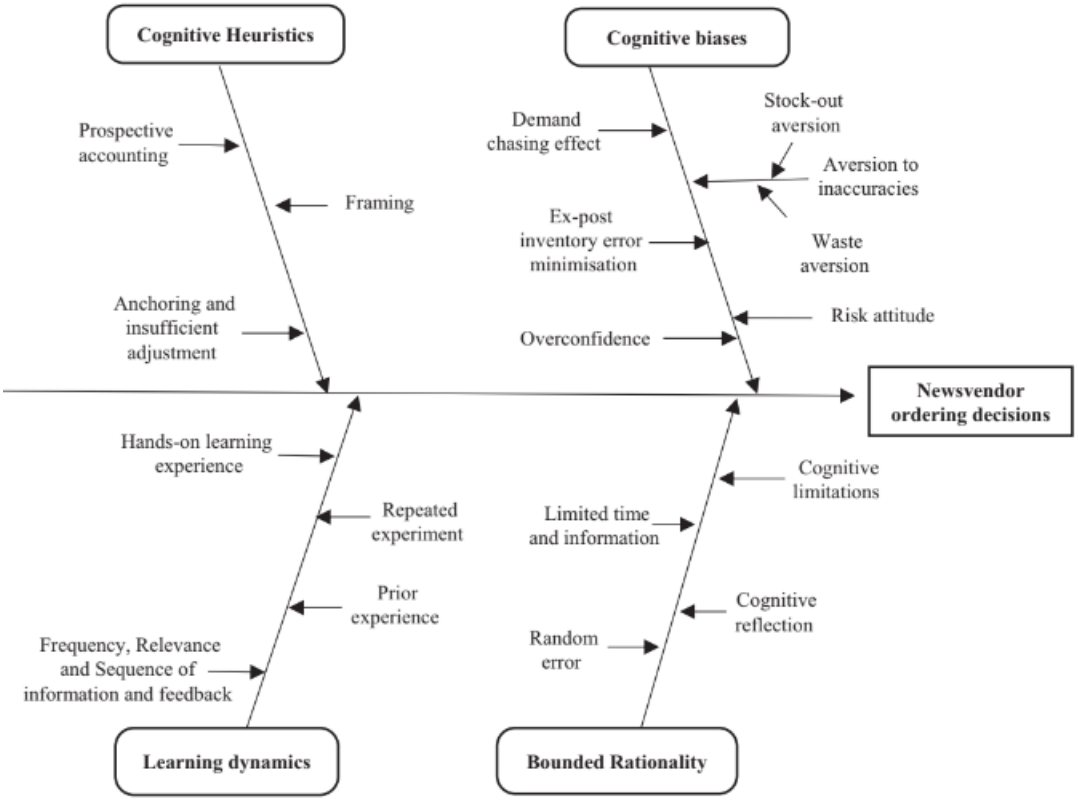

Figure 2: Elements and dynamics cooperating to the Newsvendor ordering decision

Hence, since the mathematical and the behavioural meanings have proved to be the most prominent in literature, the Newsvendor Problem can be neutrally defined, based on the mixture of the above stated definitions, as a model in which a decision maker, making his/her purchases under the stipulated conditions, produces regular deviations from optimal behaviour, which have proved to be of great interest in investigating many aspects of various generic and realistic decision making processes.

2.2. Validity of the model

A deeper investigation on the validity of the model is necessary to clarify the reasons beyond this investigation and the role it has played over the years in literature.

Starting with Bolton and Katok, on the use of the model, the authors underlined its double-face adequacy as a theoretical paradigm – given its ability to capture the relevant decision characteristics common to most inventory decisions – and as a bright test for the detection of largely applicable behavioural enlightenments [112].

Almost the same concepts are highlighted by Yamini, who wrote about the broad applicability of the model in many contexts and different decision making domains, such as inventory, revenue and pricing management, finance and supply chain contracts. Moreover, the author emphasizes the behavioural importance of the model in consideration of the fact that it reflects real scenarios in which managers must sometimes make subjective choices due to the lack of previous data and unforeseen swings in cost factors [111].

Finally, although few and very recent contributions to date, the Newsvendor Problem has proved to be relevant in the neurophysiological field too, as shown in the study conducted by Truong et al. in 2022. An analysis of decision maker neural activity and cognitive processes would lead to a deep understanding of the mechanisms underlying the decision making process [113].

Therefore, also just from the above-stated considerations, it is clear that the Newsvendor Problem relevance can be attested in many different fields, two of which were considered of particular interest for the present study and will be deepened in the following chapters: the behavioural and the neurophysiological perspectives. Hence, to assess a kind of "zero-level" relevance, the fact that the model provides a comprehensive and multifaceted understanding of the decision making process should be considered.

2.3. Approach and main contributions over time

The Newsvendor Problem has a long history dating back to the economist Edgeworth, who in 1888 introduced a variation in a bank cash-flow problem. Nevertheless, only in the 1950s the model became an object of essential studies among academicians and since then the approach with which it has been observed has gradually changed, increasingly becoming the subject of behavioural discussions (along with decision making in general). The traditional Newsvendor Problem focused on deducing the optimal order quantity that minimizes the differences between stocking too few or too many products. Instead, over the years, many started showing how inventory decisions are actually biased because of the adoption of cognitive heuristics and intuition.

In 1951 Arrow et al. modeled the first standard Newsvendor Problem, calculating the critical fractile solution explained in the previous chapter [114].

In 1955 Whitin formulated the first Newsvendor Problem with price effects, with selling price and stocking quantity set simultaneously, and a probability distribution of demand depending on the unit selling price [115].

In the same year, Simon introduced instead the concept of Bounded rationality in decision making, recognizing human intrinsic cognitive limits and deeply impacting subsequent research on the Newsvendor Problem. Until then, the theoretical observations focused on the paradigm of perfect rationality, with a perfect optimizer newsvendor [14].

2.3.1. Bullwhip effect and the Law of small number bias

An important observation that later impacted the study of the model was made by Forrester in 1958, through the formulation of the Bullwhip effect concept, which represents the propensity of orders to increase in variations advancing in a supply chain, because of different causes, such as demand signal processing and price variations [116] [117].

Still, in 1971 Tversky and Kahneman introduced instead another notion, the Law of small number bias, asserting that when taking a decision people tend to draw conclusions from improperly small samples [31].

2.3.2. Prospect theory: the importance of framing

A few years later, in 1979, the same authors formulated the Prospect theory, which raised from a critique to the expected utility theory as a descriptive paradigm of decision making under risk. The model, which deeply influenced subsequent studies, explains how decisions among risky prospects show extensive effects that are inconsistent with the expected utility theory because of a human predisposition for underestimating merely probable outcomes with respect to other certain ones, calling this fact the Certainty effect. The authors argued that this inclination promotes risk aversion in decisions concerning gains, and risk seeking in the ones concerning losses. Furthermore, the formulation of the Isolation effect notion is involved in Prospect theory since it describes how people tend to dismiss elements shared by all the possibilities under examination, leading to erratic choices when the same fact is presented in multiple ways [32].

2.3.3. Pull-to-centre effect: the importance of margin

However, probably the most famous and impacting contribution made specifically about the Newsvendor Problem was realized by Schweitzer and Cachon in 2000. The researchers conducted an experimental test and demonstrated not only that participants systematically deviate from optimal orders, but also that their order positions differ from the optimal ones depending on the product margin, specifically being above the optimal level in case of a low margin product and below the

threshold in case of a high margin one. The authors addressed Anchoring and Insufficient adjustment as one possible explanation for this phenomenon (called Pull-to-centre effect), indicating that decision-makers anchor around the average demand from the very first stages of the experiment and then insufficiently adjust in the subsequent ones toward the profit-maximizing order, failing to learn. Moreover, they also noticed that orders tended to be closer to the optimal quantity in the high margin product scenario than in the low profit one, clarifying that this pattern could not be explained either with a preference for minimizing ex-post inventory error or with anchoring and insufficient adjustment, but with stockout and waste aversion. The findings of Schweitzer and Cachon were later confirmed and hardened by many other studies [93].

2.3.4. Previous demand

In 2008, Benzion et al. discovered that orders are influenced by both the average demand and the demand of the previous period, with a decreasing influential pattern, but not sufficiently enough to make the decision maker order the optimal quantity [118].

2.3.5. Feedback and experience

Still, in 2008 Bolton and Katok extended Schweitzer and Cachon work using a hundred decision rounds, finding that the way in which feedback and experience are arranged for the decision maker may have a fundamental importance on whether inventory is stocked optimally. In detail the authors identified two main points: first, the importance of inappropriate responses inhibition to short term information, in order to keep individuals from overreactions to short term alterations; second, the fact that the greatest advancement in performance is due to gaining personal experience [112].

2.3.6. Systems 1 and 2: the importance of cognitive abilities

As another great contribution, the already mentioned publication of Kahneman's book "Thinking, Fast and Slow" in 2011 has to be considered, with its dissertation about System 1 and System 2. The work confirmed the intuitive and imperfect nature of human beings, which causes them to continuously deviate from normative predictions when making decisions [53].

2.3.7. Overconfidence

Finally, another finding of particular interest for the present study is the one that Ren and Croson realized in 2013, when they showed the importance of Overconfidence in the Newsvendor Problem.

This cognitive bias represents the decision makers tendency to underestimate the outcome of an uncertain event. The authors particularly demonstrated the positivity of the correlation between the Newsvendor ordering attitude and the bias, testing that participants can unbias their order when subjected to a specific treatment [119]. Moreover, as suggested by Moore and Healy in 2008, the bias can be distinguished in three different categories: overestimation (i.e., believing one's abilities to be superior to what they actually are), over placement (i.e., believing oneself better than others), and overprecision (i.e., believing one's estimation abilities to be superior to what they actually are) [33].

3. Research questions and hypothesis

First presented by Simon in 1955, the concept of Bounded rationality has become increasingly consolidated over time, becoming part of a broader framework representing decision makers as irrational individuals, influenced by various factors while making choices and systematically deviating from perfectly rational behaviour [14]. Over the years, many studies have indeed examined decision makers' performance to understand the dynamics underlying decision making process better, highlighting and exploring different aspects of the topic.

However, the literature review revealed a dearth of studies on one specific aspect. Indeed, while many studies focus on the biases affecting decision makers, to the best of our knowledge, a lack of researches is evident for what specifically concerns the variability inherent in the behaviour of decision makers, that is the pattern followed by individuals in deviating from rationality when deciding. The topic was thus considered of great interest for the present work, which aims to contribute to the identified literary gap.

In particular, the Newsvendor Problem was adopted for this study, considering its capability to reproduce a specific and structured scenario, which can be easily generalized and compared to various everyday life situations in which humans are affected by systematic biases.

In detail, behavioural variability was studied by analysing some possible causes of the spontaneous oscillations exhibited in the way decision makers (i.e. newsvendors) forecast demand, make orders, incur costs, and deviate from optimal values.

The analysis was thus carried out starting from the following research questions, to which the present work tried to give an answer:

1. *How much variability is generated by the experimental conditions?*
2. *How much variability is generated by the cognitive abilities of the individual?*
3. *How much variability is generated by the previous period conditions?*

From the questions, based on the available literature, the following hypotheses were hence formulated and tested:

1. *Purchase quantities are influenced by product margin* [93]
2. *Purchase quantities are influenced by framing* [120]
3. *Demand forecast is influenced by framing and not by margin* [121]

4. *Individuals learn by doing, improving performance in terms of purchases and costs distance from the optimum over sessions*
5. *The higher the individual score in the Cognitive Reflection Test, the lower the variance of the orders placed by the newsvendors* [122]
6. *The higher the academic mean of the individual, the better his/her performance in terms of purchases and costs deviation from the optimum*
7. *An individual who perceives his/her performance as belonging to the first or fourth quartile actually obtains lower performance levels in terms of purchases and costs deviation from the optimum than an individual who perceives his/her performance as belonging to the second or third quartile*
8. *The higher the variability of the demand forecast and purchase quantities, the higher the costs sustained by the newsvendors*
9. *Previous period conditions (such as previous demand and performance) influence the answering behaviour of individuals* [118]

For clarity, it is necessary to specify that the first two hypotheses (both referred to the purchase quantities) were not aggregated because of the validating nature of the first one, given the robustness of the results of the effects of margin on purchase quantities.

Finally, as pointed out in the hypotheses, in the following paragraphs, "Experimental conditions" refers to the variables framing, margin and session, "Previous period conditions" to previous values of costs or demand, and "Cognitive abilities" to the score obtained in the Cognitive Reflection Test, the participants' perceived performance, the difference between participants' perceived and actual performance, and their academic average.

4. Methods and data

The experiment in question was organized by the Politecnico di Torino in 2019 and took place with a virtual plenary session on a sample of two hundred and eighteen students from the Ingegneria Gestionale course of the University, which were divided into six groups to explore different experimental conditions. The test was scheduled in two sessions, structured as shown in Figure 3.

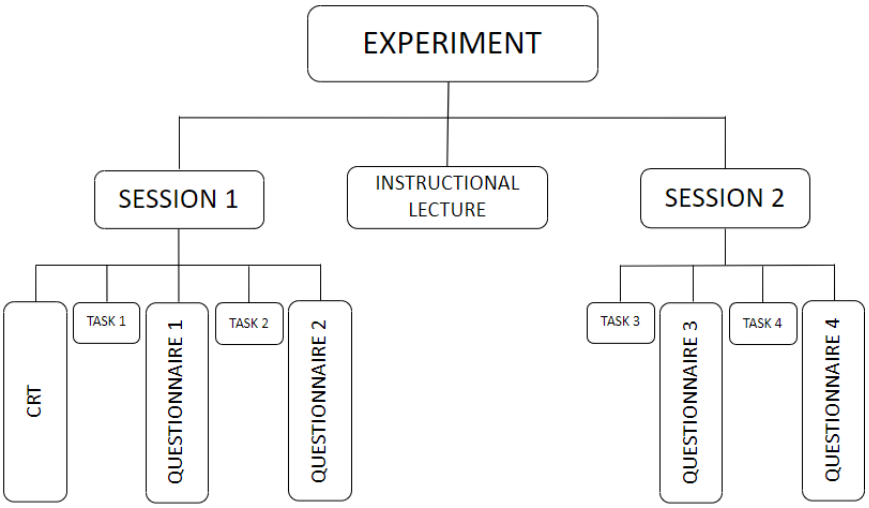

Figure 3: Structure of the experiment analysed

In total the two sessions comprised four Newsvendor Problem tasks, each followed by a questionnaire in which it was asked to participants about the strategy adopted and the perceived performance in the just finished task.

Additionally, at the beginning of the first session, the Cognitive Reflection Test (CRT) was proposed to participants, in order to test their cognitive abilities [123]. First formulated by Frederick in 2005, the CRT consists in three questions, each with an immediate but wrong answer, and a correct answer that requires instead more reflection. The test allows the classification of individuals in "Slow thinker" and "Fast thinker", resuming what stated by Kahnemann in his book "Thinking, Fast and Slow" [53], and it has already found applicability in many Newsvendor Problem experiments, such as the one conducted by Moritz et al. in 2013 [122].

An instructional lecture was instead realized between the two sessions in order to have the possibility to analyse the results of the two sessions checking whether participants had been affected by a learning effect after the lesson.

Each task of the Newsvendor Problem consisted of twenty-five rounds, with known demand for the first five periods, to see how participants behaved by already having some values from the beginning.

In every task, students had to specify both their demand forecast and purchase decisions, given that preliminarily, these two values were supposed to be both relevant for the investigation (then this supposition was also tested and proved to be right). Moreover, the tasks differed for the diverse framing-margin combinations proposed to participants, in total characterising four distinct products, namely, A-B-C-D, equal only in their 100€ unitary Purchase cost. In Table 7 it is thus reported an overview of how the experimental variables were managed during the experiment, but it has to be read in consideration of the fact that different information was given to participants, depending on the specific framing: in case they received positively framed information, they were in fact made aware of the Purchase cost, Sale price and Sellout price; instead, in case of negative framing, they were informed of the Stock-out and Overstock costs only. However, in every condition, the demand was randomly generated in vectors of length 20, characterised by an average of 100 units and a standard deviation of 30.

Table 7: Products in the Newsvendor Problem experiment analysed

PRODUCT	A	B	C	D
FRAMING	negative	negative	positive	positive
MARGIN	low	high	high	low
STOCK-OUT COST	20€	80€	80€	20€
OVERSTOCK COST	80€	20€	20€	80€
SALE PRICE	120€	180€	180€	120€
OVERSTOCK SALE PRICE	180€	120€	120€	180€

On the other hand, the differentiation in groups allowed us to propose the four products to participants in different sequences in order to explore various combinations of the experimental variables in question, as shown in Table 8. For clarity, it is however important to specify that only one hundred and forty-nine participants executed both sessions.

Table 8: Group-Task-Product combinations in the Newsvendor Problem experiment analysed

SESSION	GROUP	1	2	3	4	5	6
	TASK	PRODUCT					
1	1	A	C	A	C	A	B
	2	B	D	B	D	C	D
2	3	B	D	D	B	C	D
	4	A	C	C	A	A	B

4.1. Software and statistics adopted

The data analysis was carried out starting from an Excel dataset containing all the answers given by the participants, which was then modified and moved to the IBM SPSS software, through which all the tests and relative graphs were realized.

In detail, from a statistics perspective, the main body of the analysis consisted of many Analysis of Variance (ANOVA), which were executed to test the abovementioned influences and evaluate their statistical significance.

However, given that with respect to the costs sustained by participants the sample did not meet the assumptions of normality and homoscedasticity (essential prerequisites for the ANOVA), only to analyse the costs and their derivates, different Mann-Whitney and Kruskal-Wallis tests were conducted. Nevertheless, many graphical analysis tried to recover the main disadvantage of the Mann-Whitney and Kruskal-Wallis tests, which is that they do not allow the analysis of the impact of combined variables on a specific output, which was the element that mostly made the ANOVA suitable for the purpose of the present study.

In addition, also some Pearson correlations and T-tests were carried out to study the different variables according to the specific necessity.

4.2. Data preprocessing

Before starting the analysis, the original dataset containing 14640 participants' answers was cleaned of outliers, and so restricted to 14595 rows, according to the following upper and lower bounds: 15-200 for the purchase quantities, 25-180 for the demand forecast, -80-80 for the safety stocks, and 0-800 for the costs.

Then, to carry out the observation by task, the cleaned dataset was reduced to a shorter one, with all the necessary values calculated as averages per task, and no empty answers in terms of forecast and purchase values, thus obtaining a final dataset of 732 rows.

5. Results

The analysis was thus aimed at understanding how experimental variables, cognitive abilities and previous period conditions influence decision makers while they forecast demand, make orders, incur costs, and deviate from optimum in doing so.

The study was conducted by looking specifically at the average performance values per task, given the current lack of investigation of the setting. The original dataset values were examined only for the analysis of the previous period conditions, given the impossibility of evaluating single periods characteristics by aggregating values per task.

5.1. Validity of the sample

The first analysis realized was made to evaluate the validity of the sample and thus focused on the effect of product margin on purchase quantities, given the solidity of the literary results in this regard, as stated in the first hypothesis. To do so, a T-test was conducted to compare how participants made their purchase decisions based on the product margin.

The test showed a statistical difference in the way participants behaved in the two conditions ($p < .001$), being $\mu = 88.91$ and $\sigma = 12.23$, the values of the low margin setting, and $\mu = 106.77$ and $\sigma = 12.6$ the values of the high margin one, hence confirming that people tend to purchase differently depending on the product margin.

The results were then deepened to evaluate further the validity of what was stated by Schweitzer and Cachon in 2000, that is that people systematically deviate from profit-maximizing behaviour, with a predisposition for understocking high margin products and, vice versa, overstocking low margin products concerning the optimal order placed by the hypothetic perfectly rational newsvendor [93].

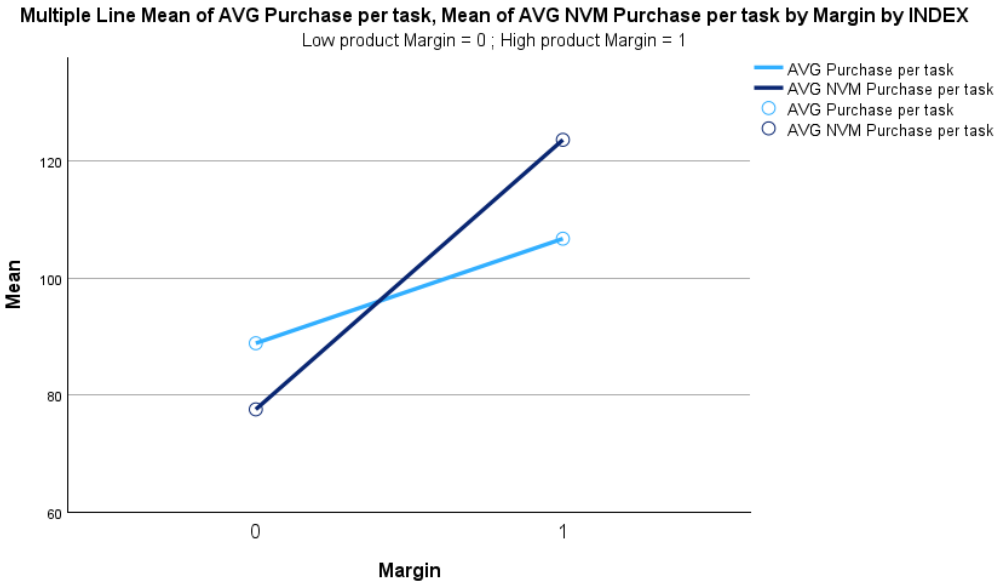

Figure 4: Line graph Margin - Purchase, Optimal Purchase

Again, literary results were confirmed, as evident in Figure 4, with participants average purchase quantities being above the optimal (NVM) level in the low margin case and below it in the high margin one.

Precisely, not only thus the averages of the actual purchase quantities found to be different by the margin, but also when compared to the optimal levels ($p < .001$), in turn, characterized by the values of $\mu = 77.61$ and $\sigma = 1.31$ for the low margin scenario, and $\mu = 123.66$ and $\sigma = 1.05$ for the high margin one.

The first hypothesis to be tested and the validity of the sample were thus confirmed.

5.2. Forecast and purchase correlation

The second step preliminary for the rest of the analysis, consisted in verifying the validity of conducting the study separately for the demand forecast and the purchase quantities.

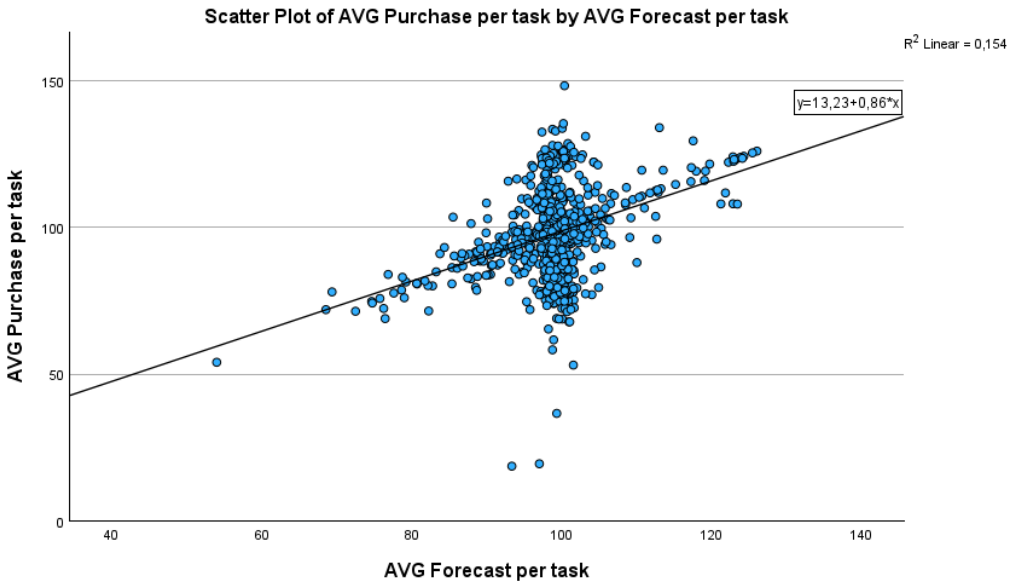

Figure 5: Pearson Correlation Forecast - Purchase

The test showed a strong positive correlation between the two variables, characterized by a Pearson correlation factor equal to 0.393, as evident in Figure 5.

However, despite the strong correlation, the two parameters were discovered to be different in their mean and standard deviation, having the purchase quantity a mean of $\mu = 97.81$ and a standard deviation of $\sigma = 15.29$, and the demand forecast with the same parameters, respectively equal to $\mu = 98.83$ and $\sigma = 7.02$.

Overall the test thus suggested that demand forecast and purchase quantities have different functions in the newsvendors' decision making process, thus supporting the idea of investigating the two factors separately.

Moreover, the test revealed another interesting result: people understood that the average correct quantity to forecast was 100, but they also varied a lot from this value when making their purchase decisions. This insight thus confirms the presence of the Pull-to-centre effect in the analysed sample (specifically at the demand forecast level), but also that people are strongly biased in their decisions [93].

5.3. Experimental variables impact

Separate ANOVAs with a 95% confidence interval and Mann-Whitney tests were then conducted in order to analyse the impact of the experimental variables (framing, margin and session) on the demand forecast, purchase quantities, costs incurred and annexed deviations from the optimum.

The first analysis was aimed at investigating the impact of these variables on both the average values per task of demand forecast and purchase quantities, through a multivariate ANOVA, as reported in Figure 6.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AVG Forecast per task	4434,016[a]	7	633,431	14,509	<,001
	AVG Purchase per task	84222,002[b]	7	12031,715	100,445	<,001
Intercept	AVG Forecast per task	6857571,273	1	6857571,273	157070,348	<,001
	AVG Purchase per task	6736479,373	1	6736479,373	56238,414	<,001
Session	AVG Forecast per task	600,644	1	600,644	13,758	<,001
	AVG Purchase per task	1062,225	1	1062,225	8,868	,003
Framing	AVG Forecast per task	128,526	1	128,526	2,944	,087
	AVG Purchase per task	3995,007	1	3995,007	33,352	<,001
Margin	AVG Forecast per task	878,267	1	878,267	20,116	<,001
	AVG Purchase per task	67287,439	1	67287,439	561,738	<,001
Session * Framing	AVG Forecast per task	1447,783	1	1447,783	33,161	<,001
	AVG Purchase per task	4668,841	1	4668,841	38,977	<,001
Session * Margin	AVG Forecast per task	1231,245	1	1231,245	28,201	<,001
	AVG Purchase per task	13411,985	1	13411,985	111,968	<,001
Framing * Margin	AVG Forecast per task	15,497	1	15,497	,355	,552
	AVG Purchase per task	523,487	1	523,487	4,370	,037
Session * Framing * Margin	AVG Forecast per task	14,767	1	14,767	,338	,561
	AVG Purchase per task	4,398	1	4,398	,037	,848
Error	AVG Forecast per task	31609,286	724	43,659		
	AVG Purchase per task	86723,838	724	119,784		
Total	AVG Forecast per task	7186112,549	732			
	AVG Purchase per task	7174253,026	732			
Corrected Total	AVG Forecast per task	36043,303	731			
	AVG Purchase per task	170945,840	731			

a. R Squared = ,123 (Adjusted R Squared = ,115)
b. R Squared = ,493 (Adjusted R Squared = ,488)

Figure 6: ANOVA Experimental variables – Forecast, Purchase

The test thus showed that while session and margin impacted both demand forecast and purchase quantities, framing only influenced the average purchase per task. This correlates with the highlighted impacts of the variables combinations, as evident in the fact that framing and margin together did not affect the demand forecast but only the purchase quantities, and suggesting that the specific session played such a fundamental role in determining the forecast output, to have been able to influence it also in relation to a specific framing. Instead, somehow surprisingly, the three variables aggregated did not significantly influence either the forecast or the purchase.

The analysis underlined the importance of the variable session, stressing the value of studying it as a driver of variability. In detail, as shown in Figure 7, on average, newsvendors elaborated higher values of both demand forecast and purchase quantities in the second task than in the first one, probably because of some acquired confidence and optimism. Moreover, the graph also allows to evidence that people tend, on average, to purchase lower quantities than the expected demand, presumably because of a strategy aimed at containing as much as possible the sellout volumes.

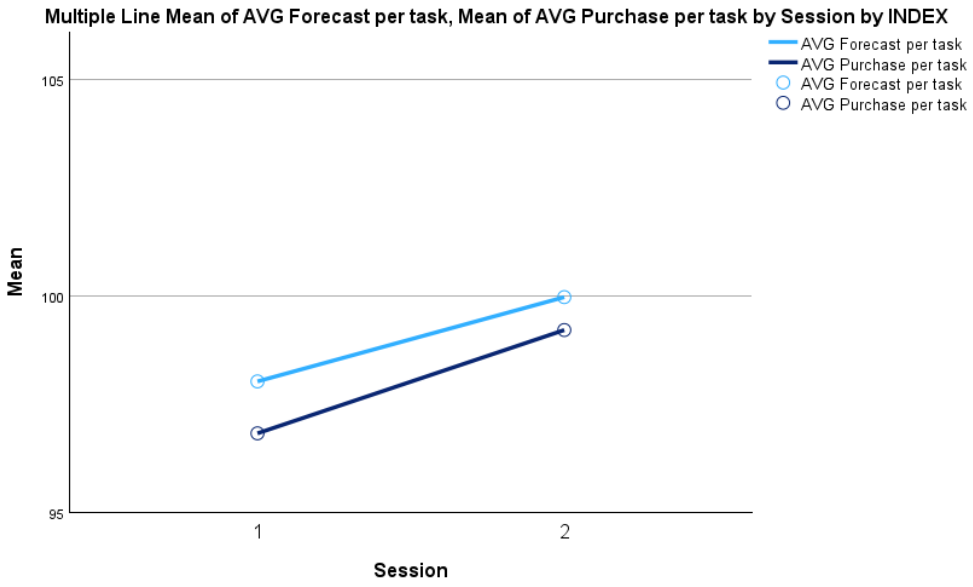

Figure 7: Line graph Session - Forecast, Purchase

The analysis also confirmed the second hypothesis, which was intended to test whether the purchase quantities are affected by framing. Indeed, as evident in Figure 8, different average quantities were ordered based on the different information framing proposed to participants. Moreover, the specific literary results from Prospect theory, about how differently newsvendors purchase depending on the framing, were also confirmed: people tend to order more products when dealing with negative framing (μ = 100.26), than with positive framing (μ = 95.59), being the two cases statistically different ($p < .001$) [32].

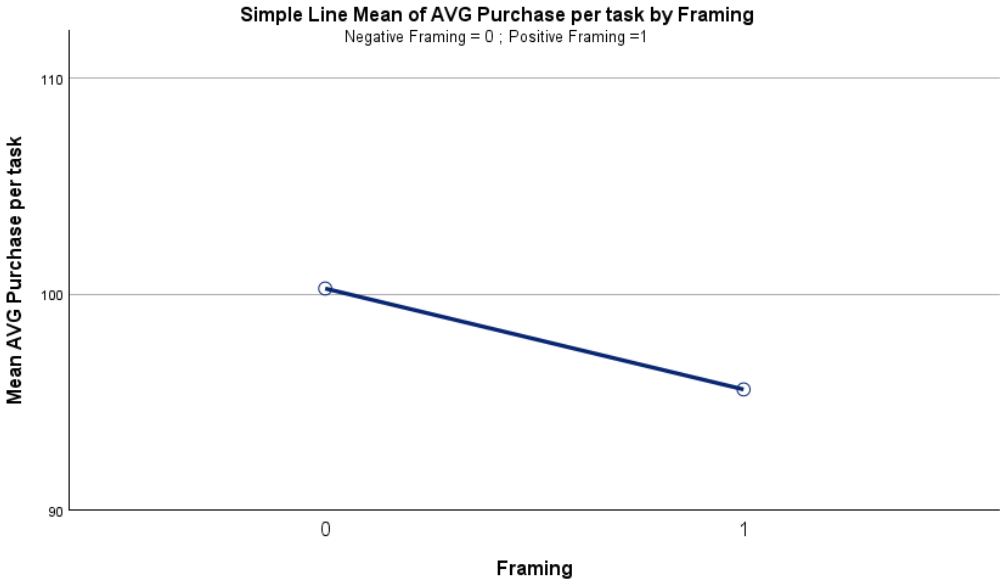

Figure 8: Line graph Framing - Purchase

The third hypothesis was instead rejected, given that the test showed the exact opposite result: at the task level, the forecast was influenced by margin and not by framing, with individuals forecasting demand statistically differently ($p < .001$), based on the profit margin of the product, in particular on average supposing lower demand values in the low margin scenario ($\mu = 97.95$) and higher values in the high margin one ($\mu = 99.72$). However, to be in contradiction with other literary research about the framing effect on demand forecast was reputed as an acceptable result overall, given the few contributions on the topic and the different points of view from which it has been investigated so far. The outcome was not considered a blocking element for the rest of the analysis.

In any case, it was ascertained that individuals tend to be deviated from optimal thinking when forecasting demand, given that an optimizing newsvendor would have forecasted a mean demand per task of 100 in every circumstance. Indeed, as evident from different T-tests, the experimental variables do not statistically impact optimal levels of both forecast and purchase quantities ($p > .05$), confirming that perfectly rational newsvendors decide independently of any contingency. As expected, however, the only condition impacting the optimizing decision maker is the margin concerning the purchase decision ($p < .001$), as correctly explained by the formula of Brandimarte and Zotteri, which represents the maximizing profit quantity $Q*$, as a function of the product cost (c) and margin ($m = p - c$) [124]:

$$F(Q*) = \frac{m}{m + c}$$

Moving on, to further analyse how the experimental conditions impact newsvendors' behaviour in terms of variability, it was also investigated how these factors influence decision makers average costs per task. Costs play indeed a fundamental role in the Newsvendor Problem in general, particularly constituting a key element in studying whether individuals learn how to improve their performance over tasks, succeeding in progressively reducing losses. However, given the nature of the study, costs were considered of great importance for the present work firstly as an output possibly characterised by significant fluctuations, because of different causes.

Thus, for this purpose, as reported in Figure 9, three separate Mann-Whitney tests were realized, each aimed at investigating the impact on costs of one specific experimental variable, reporting that session, framing and margin exercised a significant influence on the average costs per task sustained by participants. In detail, framing had a valuable effect on costs ($p = .005$), but session and margin resulted in a still more significantly impacting ($p < .001$).

Test Statistics[a]

	AVG Costs per task
Mann-Whitney U	39218,500
Wilcoxon W	84669,500
Z	-9,111
Asymp. Sig. (2-tailed)	<,001

a. Grouping Variable: Session

Test Statistics[a]

	AVG Costs per task
Mann-Whitney U	58887,500
Wilcoxon W	119962,500
Z	-2,781
Asymp. Sig. (2-tailed)	,005

a. Grouping Variable: Framing

Test Statistics[a]

	AVG Costs per task
Mann-Whitney U	49702,000
Wilcoxon W	117230,000
Z	-6,039
Asymp. Sig. (2-tailed)	<,001

a. Grouping Variable: Margin

Figure 9: Mann-Whitney test Experimental variables - Costs

These results were then deepened through a graphical analysis highlighting some interesting insights.

Particularly relevant in this regard is the representation of the impact of the session on average costs (Figure 10): individuals seem to have improved their performance in terms of costs from one session to another, being this probably attributable both to the instructional lecture and their capability of learning by doing. Hence this result underlines again the correctness of evaluating the two sessions separately.

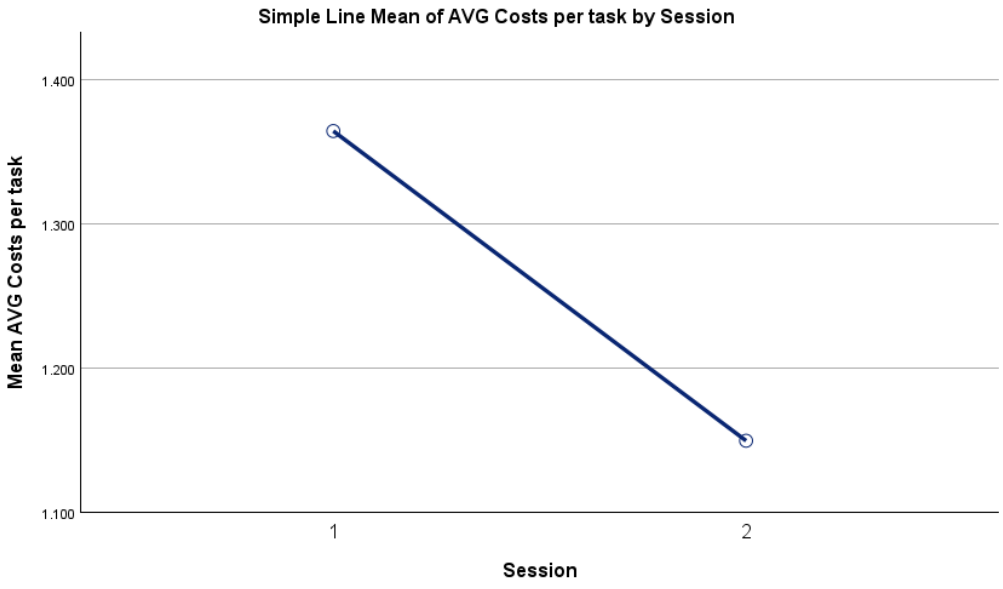

Figure 10: Line graph Session - Costs

Similar patterns characterise the graphs of framing instead (Figure 11) and margin (Figure 12), suggesting that, independently from the other conditions, people sustained lower average costs per task when dealing with negative framing (than with positive framing), and with low margin products (then with high margin ones).

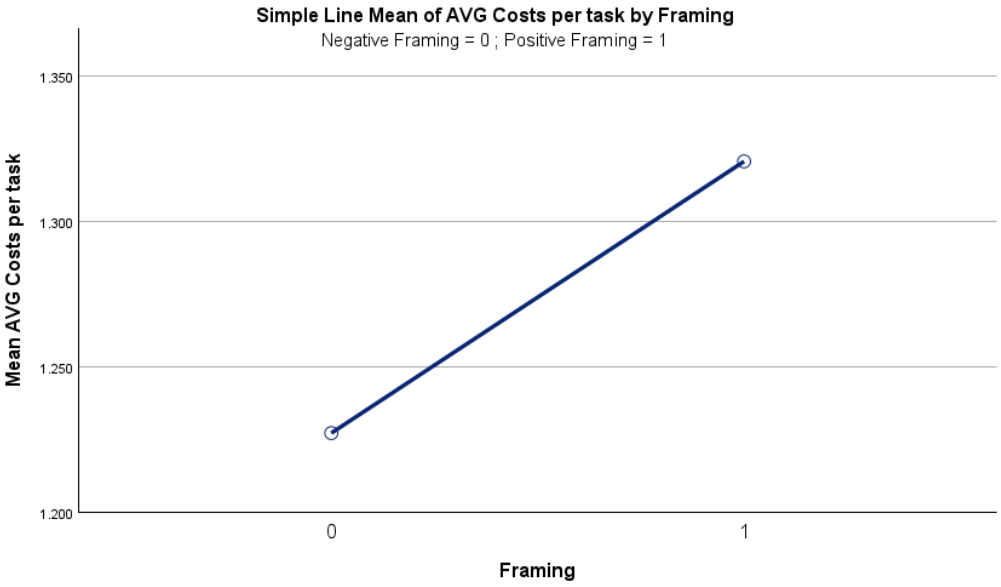

Figure 11: Line graph Framing - Costs

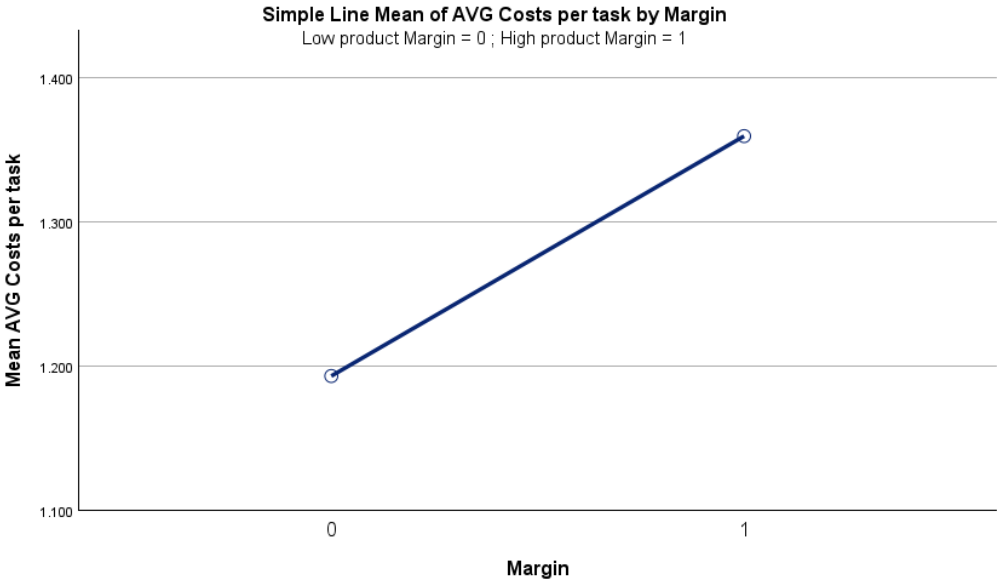

Figure 12: Line graph Margin – Costs

Moving on, to summarize the above-discussed results, the analysis investigated how framing, margin and session influence the difference between the optimal values of purchase quantities and costs, and the corresponding values given by participants. In this regard, a univariate ANOVA and a Mann-Whitney test were realized, respectively, to study the differences between purchase quantities (Figure 13) and costs (Figure 14).

Tests of Between-Subjects Effects

Dependent Variable: AVG Delta NVM Purchase per task

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	173010,334[a]	7	24715,762	206,423	<,001
Intercept	4782,444	1	4782,444	39,942	<,001
Session	657,954	1	657,954	5,495	,019
Framing	3673,208	1	3673,208	30,678	<,001
Margin	122752,906	1	122752,906	1025,219	<,001
Session * Framing	4438,741	1	4438,741	37,072	<,001
Session * Margin	14041,669	1	14041,669	117,274	<,001
Framing * Margin	1429,198	1	1429,198	11,937	<,001
Session * Framing * Margin	669,479	1	669,479	5,591	,018
Error	86686,968	724	119,733		
Total	265268,546	732			
Corrected Total	259697,302	731			

a. R Squared = ,666 (Adjusted R Squared = ,663)

Figure 13: ANOVA Experimental variables - Delta Purchase

Test Statistics[a]

	AVG Delta PurchaseCost NVMCosts per task
Mann-Whitney U	32359,000
Wilcoxon W	77810,000
Z	-11,548
Asymp. Sig. (2-tailed)	<,001

a. Grouping Variable: Session

Test Statistics[a]

	AVG Delta PurchaseCost NVMCosts per task
Mann-Whitney U	59533,500
Wilcoxon W	120608,500
Z	-2,555
Asymp. Sig. (2-tailed)	,011

a. Grouping Variable: Framing

Test Statistics[a]

	AVG Delta PurchaseCost NVMCosts per task
Mann-Whitney U	54117,500
Wilcoxon W	121645,500
Z	-4,496
Asymp. Sig. (2-tailed)	<,001

a. Grouping Variable: Margin

Figure 14: Mann-Whitney test Experimental variables - Delta Costs

The analysis confirmed and enlarged the discussed results so far: all the experimental variables played a fundamental role in determining how individuals systematically deviated from optimal values of purchase quantities and costs.

Finally, to further deepen specifically the way in which different combinations of the experimental variables originated the variability that induced participants to deviate from optimal values of purchase quantities and costs, a graphic analysis was realized in order to define the attributes of the variables mainly influencing the process, through an analysis of all the possible scenarios.

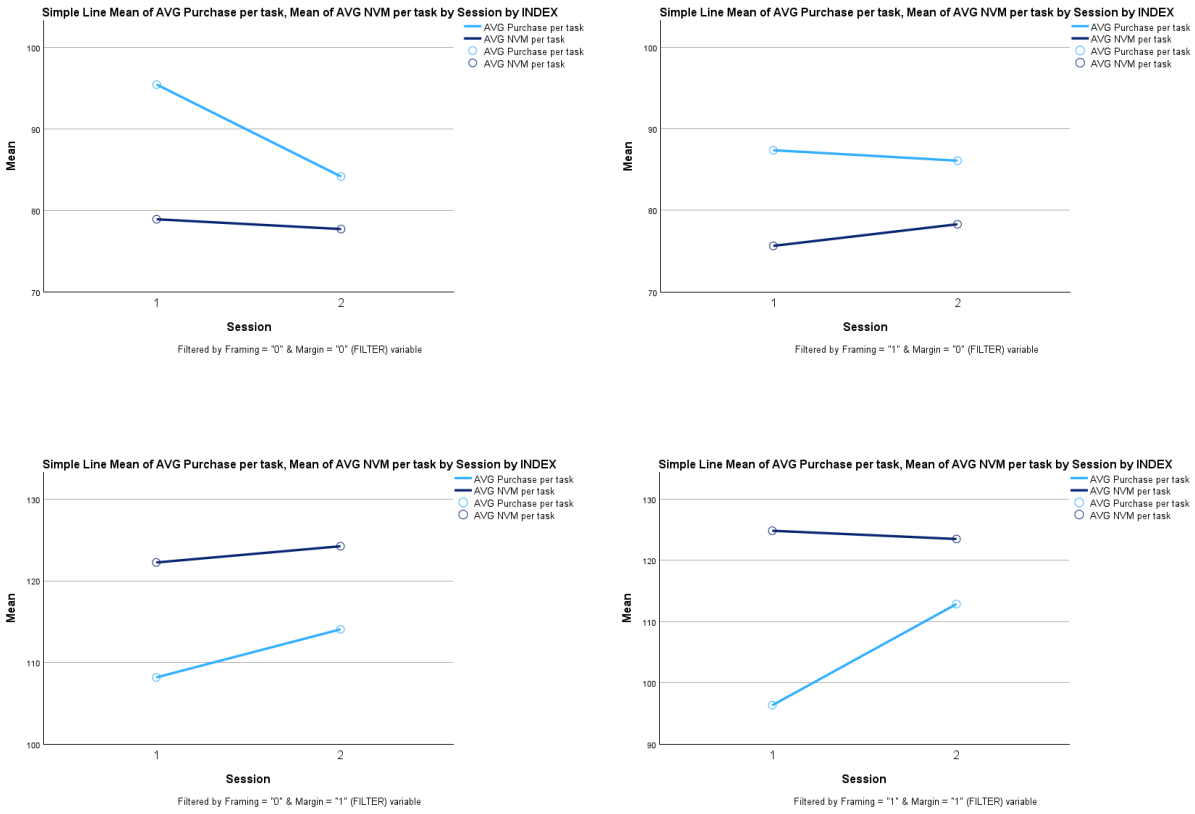

Figure 15: Line graph Combinations of Framing, Margin - Purchase, Optimal Purchase

On average, in all the possible scenarios, the deviation from the optimum was smaller in the second session than in the first one: again, this confirms the learning process undertaken by participants, or, similarly, the validity of the instructional lecture held between the two sessions, as previously observed. In addition, what was discussed above regarding the impact of margin on participants' position with respect to the optimum is also evident in the graphs: irrespectively of the other experimental conditions, participants are above the optimum in cases of low margin and vice versa below it in cases of high margin. Finally, it can be observed that the scenario characterised by an overall lower deviation from the optimum in terms of purchases was the one characterised by positive framing and low product margin. Although the average purchase per task made under these conditions by the participants ($\mu = 87.38$, $\sigma = 12.5$ for the first session; $\mu = 86.09$, $\sigma = 12.53$ for the second session) is statistically different ($p < .001$) from the average optimal value ($\mu = 75.65$, $\sigma = 0.05$ for the first session; $\mu = 78.3$, $\sigma = 0.03$ for the second session), the averages of the two indicators, for each session, differed overall less than in the other combinations.

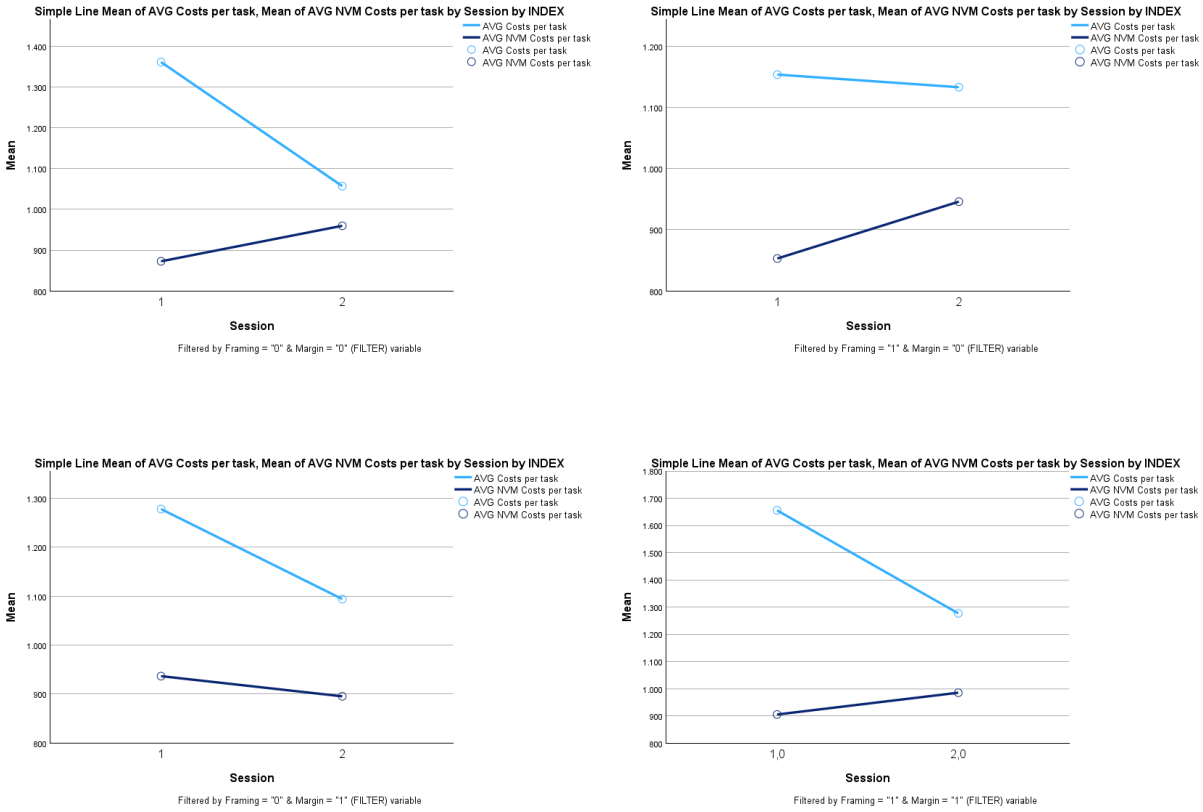

Figure 16: Line graph Combinations of Framing, Margin - Costs, Optimal Costs

Similarly, looking at the costs (Figure 16), it is evident that overall, participants managed to decrease the distance from the optimum from the first session to the subsequent one, hence improving their performance from this perspective. Moreover, in connection with what was stated above for the purchase quantities, the scenario characterized by positive framing and low product margin was the one in which participants behaved better in terms of distance from the optimum on average.

Therefore, in conclusion, it can be stated that these results confirmed the correctness of the fourth hypothesis: over sessions, individuals tend to improve in performance, becoming able to reduce the distance from the optimum both in terms of purchases and costs.

5.4. Cognitive abilities impact

The second main section of the analysis focused instead on understanding whether cognitive abilities produce variability in how newsvendors make decisions. For this purpose, again, different ANOVAs with a 95% confidence interval and Mann-Whitney tests were conducted.

In detail, a multivariate ANOVA was realized to understand if individual cognitive abilities impact demand forecast and purchase quantities, as reported in Figure 17.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	AVG Forecast per task	5739,525[a]	137	41,894	,729	,986
	AVG Purchase per task	21483,713[b]	137	156,815	,550	1,000
Intercept	AVG Forecast per task	2656104,368	1	2656104,368	46200,052	<,001
	AVG Purchase per task	2535213,221	1	2535213,221	8891,230	<,001
AverageExams	AVG Forecast per task	3149,816	82	38,412	,668	,987
	AVG Purchase per task	12043,917	82	146,877	,515	1,000
CRTCorrect	AVG Forecast per task	72,407	3	24,136	,420	,739
	AVG Purchase per task	174,208	3	58,069	,204	,894
PerceivedPerformance	AVG Forecast per task	9,249	2	4,624	,080	,923
	AVG Purchase per task	66,625	2	33,312	,117	,890
PercVSRealPerformance	AVG Forecast per task	149,473	4	37,368	,650	,627
	AVG Purchase per task	314,421	4	78,605	,276	,894
AverageExams * CRTCorrect	AVG Forecast per task	441,762	4	110,440	1,921	,106
	AVG Purchase per task	576,012	4	144,003	,505	,732
AverageExams * PerceivedPerformance	AVG Forecast per task	,000	0	.	.	.
	AVG Purchase per task	,000	0	.	.	.
AverageExams * PercVSRealPerformance	AVG Forecast per task	110,811	3	36,937	,642	,588
	AVG Purchase per task	110,263	3	36,754	,129	,943
CRTCorrect * PerceivedPerformance	AVG Forecast per task	,000	0	.	.	.
	AVG Purchase per task	,000	0	.	.	.
CRTCorrect * PercVSRealPerformance	AVG Forecast per task	262,624	4	65,656	1,142	,336
	AVG Purchase per task	1078,973	4	269,743	,946	,437
PerceivedPerformance * PercVSRealPerformance	AVG Forecast per task	,000	0	.	.	.
	AVG Purchase per task	,000	0	.	.	.
AverageExams * CRTCorrect * PerceivedPerformance	AVG Forecast per task	,000	0	.	.	.
	AVG Purchase per task	,000	0	.	.	.
AverageExams * CRTCorrect * PercVSRealPerformance	AVG Forecast per task	,000	0	.	.	.
	AVG Purchase per task	,000	0	.	.	.
AverageExams * PerceivedPerformance * PercVSRealPerformance	AVG Forecast per task	,000	0	.	.	.
	AVG Purchase per task	,000	0	.	.	.
CRTCorrect * PerceivedPerformance * PercVSRealPerformance	AVG Forecast per task	,000	0	.	.	.
	AVG Purchase per task	,000	0	.	.	.
AverageExams * CRTCorrect * PerceivedPerformance * PercVSRealPerformance	AVG Forecast per task	,000	0	.	.	.
	AVG Purchase per task	,000	0	.	.	.
Error	AVG Forecast per task	25123,730	437	57,491		
	AVG Purchase per task	124604,601	437	285,136		
Total	AVG Forecast per task	5673598,385	575			
	AVG Purchase per task	5662587,797	575			
Corrected Total	AVG Forecast per task	30863,254	574			
	AVG Purchase per task	146088,314	574			

a. R Squared = ,186 (Adjusted R Squared = -,069)

b. R Squared = ,147 (Adjusted R Squared = -,120)

Figure 17: ANOVA Cognitive abilities - Forecast, Purchase

Quite surprisingly, the test reported that at a task level, in the examined sample, cognitive abilities did not significantly influence newsvendors either in their demand forecast or in their purchase decisions, being all the significances reported in the table higher than .05.

This result thus opened the way for the examination of the fifth, sixth and seventh hypotheses, which concern specifically the variability of the orders and the distance from the optimum, here still not taken into account.

Hence, to deepen the influence of the cognitive abilities on the Newsvendor Problem, another analysis was then conducted in order to understand whether these influenced the costs sustained by participants. To this end, a Kruskal-Wallis test and three separate Mann-Whitney tests were produced, respectively to examine the influence of participants' perceived performance on the average costs (Figure 18), and of the impact of the other three examined cognitive abilities (Figure 19).

However, it is necessary to specify that these last three cognitive abilities were transformed in their binary form to carry out the investigation, given that the Mann-Whitney test requires independent variables with only two groups of possible realizations. Instead, for what concerns perceived performance, this was not purposely modified since its investigation aimed to open the way for the test of the seventh hypothesis. In detail, the modification applied to the variables in question entailed the following changes: the number of correct answers given to the CRT was replaced with a zero in case of Slow thinker (i.e., three correct answers) or with a one in case of Fast thinker (i.e. less than three correct answers); for the academic average, twenty-five was assumed as threshold distinguishing high academic paths (one) and low ones (zero); finally, the difference in the performance perception were attributed to a zero in case of value lower than zero, or vice versa to a one in case of values equal or higher than zero.

Test Statistics[a,b]

	AVG Costs per task
Kruskal-Wallis H	,963
df	3
Asymp. Sig.	,810

a. Kruskal Wallis Test

b. Grouping Variable: Perceived Performance

Figure 18: Kruskal-Wallis test Perceived Performance - Costs

Test Statistics[a]

	AVG Costs per task
Mann-Whitney U	33884,500
Wilcoxon W	43754,500
Z	-3,358
Asymp. Sig. (2-tailed)	<,001

a. Grouping Variable: Binary_CRTCorrect

Test Statistics[a]

	AVG Costs per task
Mann-Whitney U	14908,500
Wilcoxon W	114589,500
Z	-8,339
Asymp. Sig. (2-tailed)	<,001

a. Grouping Variable: Binary_PercVSRealPerformance

Test Statistics[a]

	AVG Costs per task
Mann-Whitney U	41787,000
Wilcoxon W	56152,000
Z	-2,400
Asymp. Sig. (2-tailed)	,016

a. Grouping Variable: Binary_AverageExams

Figure 19: Mann-Whitney test Cognitive abilities - Costs

In detail, the test showed that perceived performance did not have a relevant role in determining the average costs sustained by participants (significance level > .05), differently from the other three examined participants' cognitive abilities which (even if in their approximated binary form) had a significant influence, specifically for what concerns the correct answers given to the CRT and the performance perception indicator ($p < .001$).

In detail, in its not-binary form, the difference between the perceived and the actual participants' performance must be observed from two different perspectives. First, it indeed suggests the correctness of the prediction made by the participant, in the sense that the closer its amount to zero, the higher the cognitive abilities of the candidate, given that it would mean that he/she was able to correctly identify the quartile in which his/her performance was actually placed. On the other hand, a greater value of the indicator also suggests an actual better performance, despite the correctness of the prediction.

These two points of view clarify what is visible in Figure 20: as supposed, the pattern of the average costs is declining with the performance indicator, in detail showing that the good performance entailed in a higher value of this is more relevant than the correctness of the prediction suggested by a value near to zero, with the view of lowering costs. Moreover, the graphic representation points out also that a wider variability range characterised by lower values of the indicator in terms of average costs per task, hence proposing that overall the higher this type of cognitive ability, the lower the cost variability.

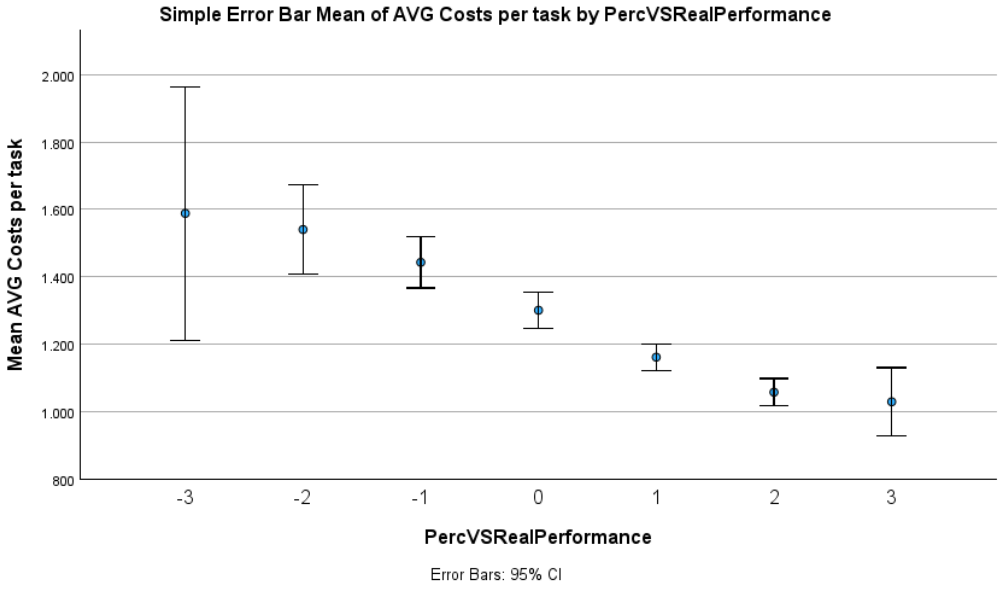

Figure 20: Simple Error Bar Delta Performance – Costs

A similar decreasing pattern also characterizes the other mainly impacting cognitive ability, that is, the number of correct answers given to the CRT, as visible in Figure 21, showing again the positive influence of higher cognitive abilities in lowering costs.

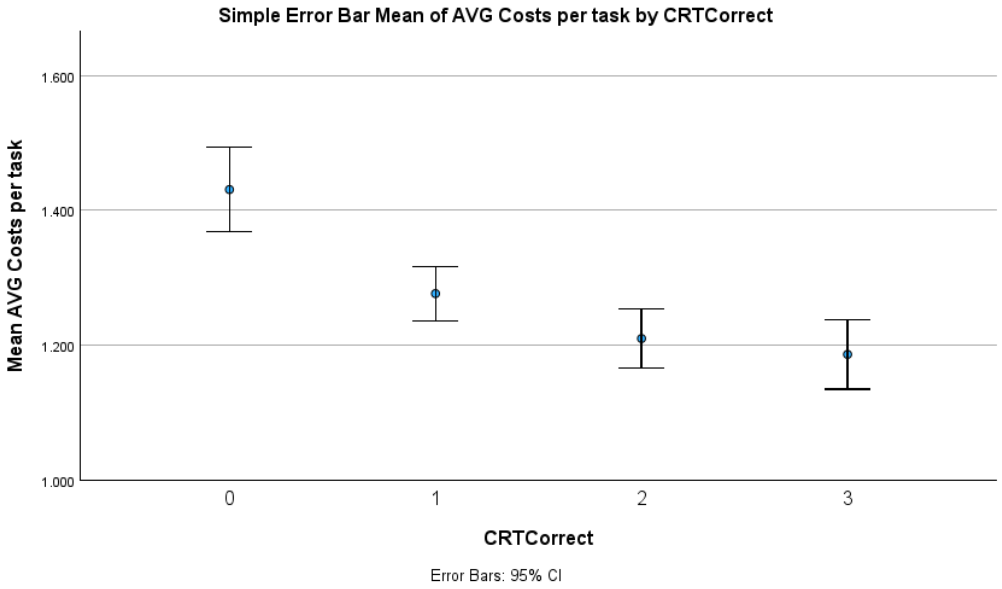

Figure 21: Simple Error Bar CRT Correct – Costs

In conclusion, another analysis was realized to further explore the discussed results to understand if cognitive abilities impact the deviations from the optimal values of purchase quantities and costs.

In this regard, a univariate ANOVA, a Kruskal-Wallis test and three other Mann-Whitney tests were conducted, resulting in the confirmation of the already mentioned results: while participants' cognitive abilities did not affect the distance of the purchase quantities from the optimum, they influenced the distance from the optimum in terms of costs, with the only exception of the perceived performance per se.

Tests of Between-Subjects Effects

Dependent Variable: AVG Delta NVM Purchase per task

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	21652,814[a]	137	158,050	,424	1,000
Intercept	3773,772	1	3773,772	10,118	,002
AverageExams	12147,441	82	148,140	,397	1,000
CRTCorrect	174,319	3	58,106	,156	,926
PerceivedPerformance	48,295	2	24,148	,065	,937
PercVSRealPerformance	340,492	4	85,123	,228	,923
AverageExams * CRTCorrect	572,183	4	143,046	,384	,820
AverageExams * PerceivedPerformance	,000	0	.	.	.
AverageExams * PercVSRealPerformance	98,891	3	32,964	,088	,966
CRTCorrect * PerceivedPerformance	,000	0	.	.	.
CRTCorrect * PercVSRealPerformance	1070,256	4	267,564	,717	,580
PerceivedPerformance * PercVSRealPerformance	,000	0	.	.	.
AverageExams * CRTCorrect * PerceivedPerformance	,000	0	.	.	.
AverageExams * CRTCorrect * PercVSRealPerformance	,000	0	.	.	.
AverageExams * PerceivedPerformance * PercVSRealPerformance	,000	0	.	.	.
CRTCorrect * PerceivedPerformance * PercVSRealPerformance	,000	0	.	.	.
AverageExams * CRTCorrect * PerceivedPerformance * PercVSRealPerformance	,000	0	.	.	.
Error	162995,570	437	372,988		
Total	188901,875	575			
Corrected Total	184648,384	574			

a. R Squared = ,117 (Adjusted R Squared = -,159)

Figure 22: ANOVA Cognitive abilities - Delta Purchase

Hence the test made evidence of the rejection of the sixth and seventh hypotheses for what concerns specifically the purchase deviation from the optimum: indeed, this was not significantly impacted either by the academic average, nor by the perceived performance ($p > .05$).

Test Statistics[a,b]

	AVG Delta PurchaseCost NVMCosts per task
Kruskal-Wallis H	1,382
df	3
Asymp. Sig.	,710

a. Kruskal Wallis Test

b. Grouping Variable: Perceived Performance

Figure 23: Kruskal-Wallis test Perceived Performance - Delta Costs

Test Statistics[a]

	AVG Delta PurchaseCost NVMCosts per task
Mann-Whitney U	41767,500
Wilcoxon W	56132,500
Z	-2,408
Asymp. Sig. (2-tailed)	,016

a. Grouping Variable: Binary_AverageExams

Test Statistics[a]

	AVG Delta PurchaseCost NVMCosts per task
Mann-Whitney U	33851,500
Wilcoxon W	43721,500
Z	-3,373
Asymp. Sig. (2-tailed)	<,001

a. Grouping Variable: Binary_CRTCorrect

Test Statistics[a]

	AVG Delta PurchaseCost NVMCosts per task
Mann-Whitney U	15236,000
Wilcoxon W	114917,000
Z	-8,142
Asymp. Sig. (2-tailed)	<,001

a. Grouping Variable: Binary_PercVSRealPerformance

Figure 24: Mann-Whitney test Cognitive abilities – Delta Costs

Instead, different considerations have to be made about the two hypotheses (i.e. the sixth and the seventh) concerning the distance from the optimal costs. Indeed, while the Kruskal-Wallis test showed that the perceived performance did not have a valuable influence, the Mann-Whitney tests revealed that the other three cognitive abilities played a relevant role in determining the deviations from the optimal costs.

These results were then deepened graphically for what concerns the perceived performance impact, despite its lack of statistical relevance. Hence, two hypothesis-specific graphical representations were realized (Figure 25 and 26), respectively, using the binary form of the academic performance (because of a better depiction with respect to the not-transformed variable) and the pure perceived performance (because of the specific content to be tested).

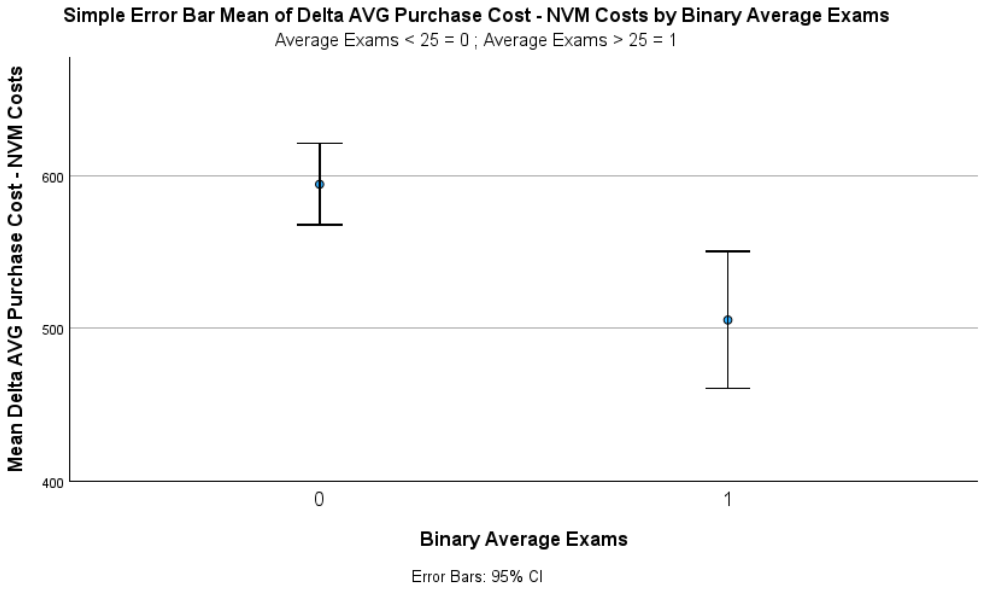

Figure 25: Simple Error Bar Binary Average Exams - Delta Costs

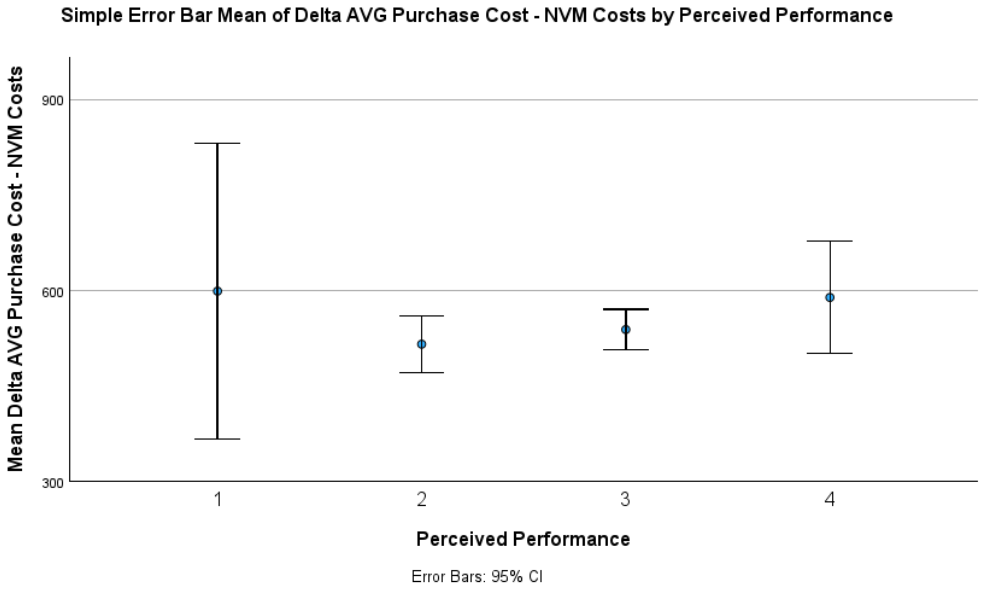

Figure 26: Simple Error Bar Perceived Performance - Delta Costs

The two graphs thus show how the participants with a higher academic mean and who perceived themselves as belonging to the second or third quartile on average deviated less from the optimal costs than the other students. In detail, the result of Figure 26 thus shows that a "medium" answering behaviour with respect to the perceived performance represents a form of cognitive ability and also underlines how variability represents a negative feature of the answering behaviour. Instead, with respect to the academic mean, quite surprisingly, higher variability characterised better performance levels.

Thus, in conclusion, the sixth hypothesis was accepted concerning the costs distance from the optimum, but not the purchase distance from it. On the other hand, the seventh hypothesis was overall rejected because of the lack of a relevant statistical influence on either of the two distances, but it was graphically assessed that a pattern reflecting the hypothesis effectively exists.

5.5. Forecast and Purchase variability impact

Another possible cause of behavioural variability that was considered of fundamental importance to be investigated in the so far discussed Newsvendor Problem context is the interaction between the variability characterising the values of demand forecast and purchase quantities and the performance achieved by participants in terms of costs.

This part of the study thus aims at constituting a sort of link among the up to now emerged evidences, in detail observing the standard deviation values of demand forecast and purchase quantities, through the implementation of a Kruskal-Wallis test.

Indeed, as statistical analysis, given the already mentioned inadequacy of the ANOVA for the study of costs, it was necessary to find another test, but in this case, the use of the Mann-Whitney one was not deemed appropriate, given that it would have implied the conversion of the independent variables (i.e. the standard deviations) in binary values. The Kruskal-Wallis test was instead found to be useful, even if it does not allow either (as the Mann-Whitney test) the simultaneous comparison of different variables.

However, before starting, a short analysis was conducted specifically on the standard deviation values of forecast and purchases, independently from the costs.

In detail, in this phase, firstly, the correlation between forecast and purchase, observed at the beginning of the analysis, was reviewed graphically in terms of standard deviations. The result (Figure 27) confirmed that the two values are highly correlated, almost depicting a perfectly linear relationship.

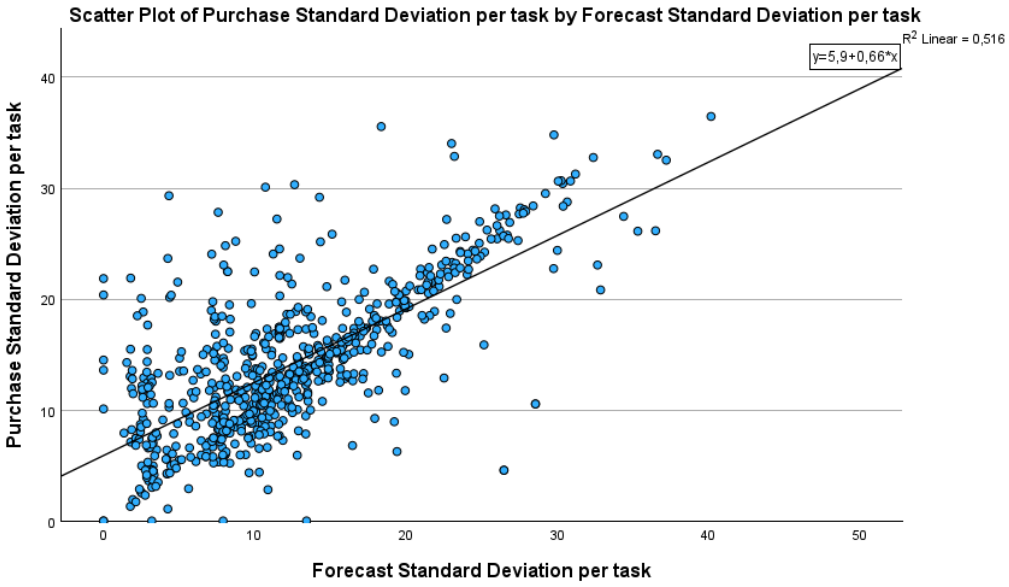

Figure 27: Scatter Plot Forecast Standard Deviation - Purchase Standard Deviation

Then, it was also analysed whether the two standard deviations were impacted by the so far discussed values of experimental variables (Figure 28) and participants' cognitive abilities (Figure 29), through two multivariate ANOVAs, respectively dedicated to the two groups of independent variables.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	StandardDeviationForecast_per_task	539,990[a]	7	77,141	1,521	,157
	StandardDeviationPurchase_per_task	964,692[b]	7	137,813	3,258	,002
Intercept	StandardDeviationForecast_per_task	102681,131	1	102681,131	2024,653	<,001
	StandardDeviationPurchase_per_task	133278,604	1	133278,604	3150,532	<,001
Session	StandardDeviationForecast_per_task	67,876	1	67,876	1,338	,248
	StandardDeviationPurchase_per_task	316,667	1	316,667	7,486	,006
Framing	StandardDeviationForecast_per_task	45,153	1	45,153	,890	,346
	StandardDeviationPurchase_per_task	69,181	1	69,181	1,635	,201
Margin	StandardDeviationForecast_per_task	,003	1	,003	,000	,994
	StandardDeviationPurchase_per_task	140,200	1	140,200	3,314	,069
Session * Framing	StandardDeviationForecast_per_task	49,861	1	49,861	,983	,322
	StandardDeviationPurchase_per_task	157,054	1	157,054	3,713	,054
Session * Margin	StandardDeviationForecast_per_task	,589	1	,589	,012	,914
	StandardDeviationPurchase_per_task	8,237	1	8,237	,195	,659
Framing * Margin	StandardDeviationForecast_per_task	336,713	1	336,713	6,639	,010
	StandardDeviationPurchase_per_task	176,427	1	176,427	4,170	,041
Session * Framing * Margin	StandardDeviationForecast_per_task	7,385	1	7,385	,146	,703
	StandardDeviationPurchase_per_task	18,536	1	18,536	,438	,508
Error	StandardDeviationForecast_per_task	36717,961	724	50,715		
	StandardDeviationPurchase_per_task	30627,749	724	42,304		
Total	StandardDeviationForecast_per_task	145506,464	732			
	StandardDeviationPurchase_per_task	173824,158	732			
Corrected Total	StandardDeviationForecast_per_task	37257,951	731			
	StandardDeviationPurchase_per_task	31592,442	731			

a. R Squared = ,014 (Adjusted R Squared = ,005)

b. R Squared = ,031 (Adjusted R Squared = ,021)

Figure 28: ANOVA Experimental variables – Forecast, Purchase Standard Deviation

As evident, the results suggested that while the session played an important role ($p < .001$) in determining the variability of the purchase quantities only, specific combinations of framing and margin impacted demand forecast and purchase decisions instead.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	StandardDeviationForecast_per_task	18238,831[a]	137	133,130	6,370	<,001
	StandardDeviationPurchase_per_task	15081,995[b]	137	110,088	5,336	<,001
Intercept	StandardDeviationForecast_per_task	40224,865	1	40224,865	1924,722	<,001
	StandardDeviationPurchase_per_task	51026,714	1	51026,714	2473,251	<,001
AverageExams	StandardDeviationForecast_per_task	10384,904	82	126,645	6,060	<,001
	StandardDeviationPurchase_per_task	6390,123	82	77,928	3,777	<,001
CRTCorrect	StandardDeviationForecast_per_task	2,295	3	,765	,037	,991
	StandardDeviationPurchase_per_task	73,273	3	24,424	1,184	,315
PerceivedPerformance	StandardDeviationForecast_per_task	453,796	2	226,898	10,857	<,001
	StandardDeviationPurchase_per_task	387,909	2	193,954	9,401	<,001
PercVSRealPerformance	StandardDeviationForecast_per_task	146,480	4	36,620	1,752	,138
	StandardDeviationPurchase_per_task	228,805	4	57,201	2,773	,027
AverageExams * CRTCorrect	StandardDeviationForecast_per_task	245,113	4	61,278	2,932	,021
	StandardDeviationPurchase_per_task	78,341	4	19,585	,949	,435
AverageExams * PerceivedPerformance	StandardDeviationForecast_per_task	,000	0	.	.	.
	StandardDeviationPurchase_per_task	,000	0	.	.	.
AverageExams * PercVSRealPerformance	StandardDeviationForecast_per_task	640,330	3	213,443	10,213	<,001
	StandardDeviationPurchase_per_task	480,452	3	160,151	7,762	<,001
CRTCorrect * PerceivedPerformance	StandardDeviationForecast_per_task	,000	0	.	.	.
	StandardDeviationPurchase_per_task	,000	0	.	.	.
CRTCorrect * PercVSRealPerformance	StandardDeviationForecast_per_task	312,897	4	78,224	3,743	,005
	StandardDeviationPurchase_per_task	399,346	4	99,836	4,839	<,001
PerceivedPerformance * PercVSRealPerformance	StandardDeviationForecast_per_task	,000	0	.	.	.
	StandardDeviationPurchase_per_task	,000	0	.	.	.
AverageExams * CRTCorrect * PerceivedPerformance	StandardDeviationForecast_per_task	,000	0	.	.	.
	StandardDeviationPurchase_per_task	,000	0	.	.	.
AverageExams * CRTCorrect * PercVSRealPerformance	StandardDeviationForecast_per_task	,000	0	.	.	.
	StandardDeviationPurchase_per_task	,000	0	.	.	.
AverageExams * PerceivedPerformance * PercVSRealPerformance	StandardDeviationForecast_per_task	,000	0	.	.	.
	StandardDeviationPurchase_per_task	,000	0	.	.	.
CRTCorrect * PerceivedPerformance * PercVSRealPerformance	StandardDeviationForecast_per_task	,000	0	.	.	.
	StandardDeviationPurchase_per_task	,000	0	.	.	.
AverageExams * CRTCorrect * PerceivedPerformance * PercVSRealPerformance	StandardDeviationForecast_per_task	,000	0	.	.	.
	StandardDeviationPurchase_per_task	,000	0	.	.	.
Error	StandardDeviationForecast_per_task	9132,888	437	20,899		
	StandardDeviationPurchase_per_task	9015,937	437	20,631		
Total	StandardDeviationForecast_per_task	109833,442	575			
	StandardDeviationPurchase_per_task	132147,010	575			
Corrected Total	StandardDeviationForecast_per_task	27371,718	574			
	StandardDeviationPurchase_per_task	24097,932	574			

a. R Squared = ,666 (Adjusted R Squared = ,562)

b. R Squared = ,626 (Adjusted R Squared = ,509)

Figure 29: ANOVA Cognitive abilities – Forecast, Purchase Standard Deviation

On the other hand, with respect to cognitive abilities, among all the possible single elements, only the academic mean and the CRT score did not have a specific impact, respectively on none of the investigated standard deviations and on the forecast one. Conversely, only some combinations of the variables had statistical relevance in the investigation (significance level < .05)

Following what was stated in the previous chapter, this test proved the fifth hypothesis to be partially correct. Specifically in terms of standard deviation (hence of variability), the score obtained in the CRT influences the variability of the orders placed by newsvendors, but only when combined with specific realizations of other forms of cognitive ability.

Moreover, even if the relevance of the CRT score per se did not have a statistically relevant influence on the test, by diagramming it with respect to the purchase quantities standard deviation, a decreasing pattern of the variance per task with the CRT can be observed, underlying the importance of a good CRT score in terms of low variability.

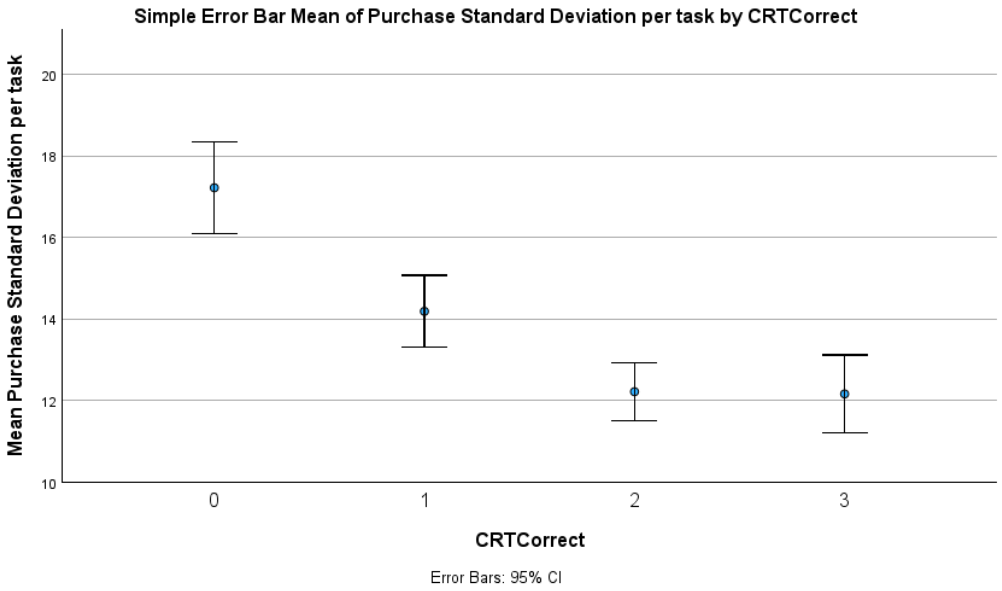

Figure 30: Simple Error Bar CRT Correct - Purchase Standard Deviation

These results thus enlarged the influence of the experimental variables and cognitive abilities on demand forecast and purchase quantities, but pointed out that diverse elements characterize the realization of the average values and, on the other hand, its variability across tasks.

Finally, after assessing what influences the two values, these were observed in relation to the average costs.

Test Statistics[a,b]

	AVG Costs per task
Kruskal-Wallis H	94,655
df	37
Asymp. Sig.	<,001

a. Kruskal Wallis Test

b. Grouping Variable: Forecast Standard Deviation per task

Test Statistics[a,b]

	AVG Costs per task
Kruskal-Wallis H	123,108
df	36
Asymp. Sig.	<,001

a. Kruskal Wallis Test

b. Grouping Variable: Purchase Standard Deviation per task

Figure 31: Kruskal-Wallis test Forecast, Purchase Standard Deviation - Costs

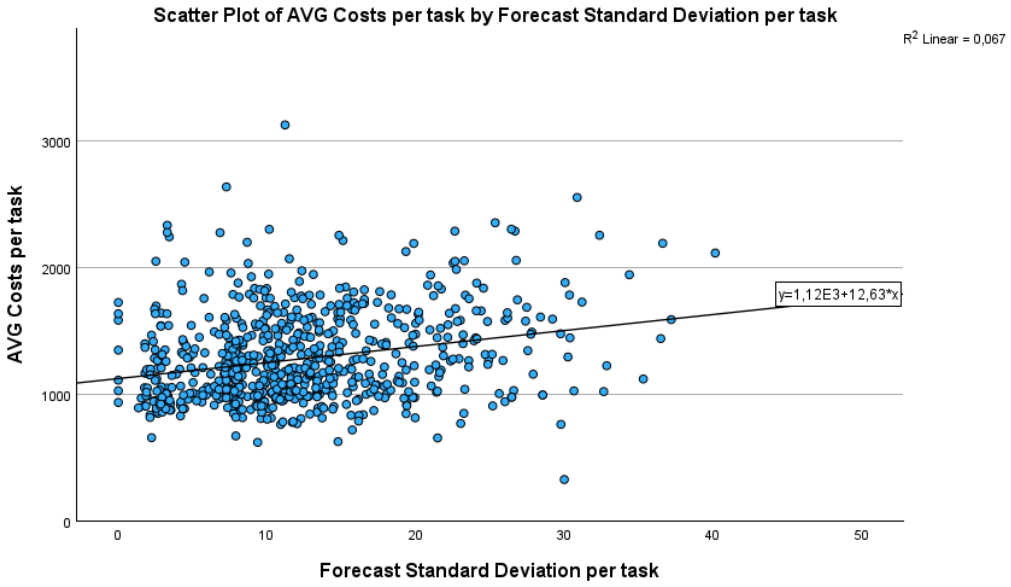

Figure 32: Scatter Plot Forecast Standard Deviation - Costs

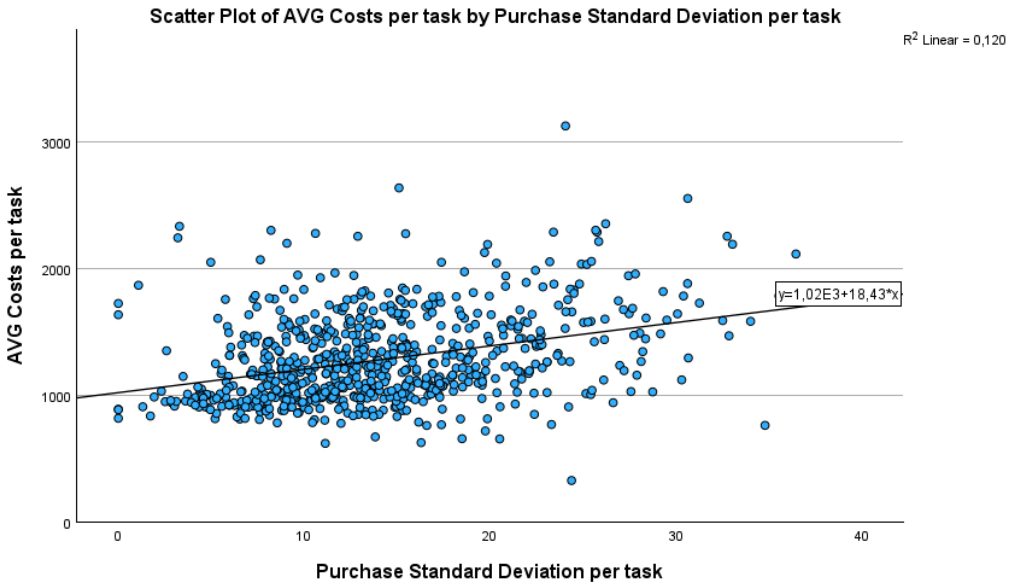

Figure 33: Scatter Plot Purchase Standard Deviation – Costs

As it can be observed both analytically (Figure 31) and graphically (Figure 32, 33), the standard deviations of demand forecast and purchases significantly influenced the average costs sustained by participants per task (significance level < .001). Moreover, specifically in the graphs, it can also be observed how these values influenced costs: on average, it is evident that costs tend to increase with the standard deviation of both forecast and purchase.

This result hence confirmed the detrimental role played by variability in the Newsvendor Problem and the acceptance of the eighth hypothesis, suggesting that the higher the dispersion of the demand forecast and purchase quantities values, the higher the average costs sustained by participants per task, thus the worst their performance.

5.6. Previous period conditions impact

Finally, the very last investigation conducted for the purpose of the present dissertation focused on the way in which previous period conditions influence decision makers in executing the Newsvendor Problem.

Differently from what was viewed in the previous chapters, this conclusive analysis was realized not by looking at the participants' averages per task but their single-period values, given the lack of reasoning in investigating how previous tasks influenced the subsequent ones.

Moreover, dissimilarly from the other analysis, it is admitted that this last test was conducted non-exhaustively. Indeed, despite the huge variety of existing period conditions, a focus was realized only on the way participants incur in cost subsequently to a shock (or a missing shock) in the previous period demand. In detail, this topic was considered of great relevance for the present paper, since it represents an important possible source of variability in the answering behaviour of decision makers. However, given the theme's amplitude, it was impossible to assess all the existing influences.

Specifically, the realization of a shock in the previous demand was considered with respect to the threshold μ +/- σ, attributing a 0 in case of demand permanence in the identified interval or a -1/1, respectively, in case of negative and positive shock.

The conducted analysis hence consisted of evaluating the Pearson correlation among the two factors, which resulted in a negative influence value equal to -0.5, suggesting that people decrease costs as the positiveness of the shock increases.

Moreover, by diagramming the two values, two other interesting results are evident: first, despite the overall negative correlation, the highest costs were sustained in case of no shock, probably suggesting that people are more prone to modify their conduct when they are subjected to a strong modification in the external (demand) values; on the other hand, the fact that the highest variability levels were registered in case of shock points out how the variety of individual behaviours also determines how people react.

Thus the test proved the ninth hypothesis correct: even if it was ascertained only with respect to the shocks in previous period demand, precedent period conditions seem to influence the behaviour of decision makers.

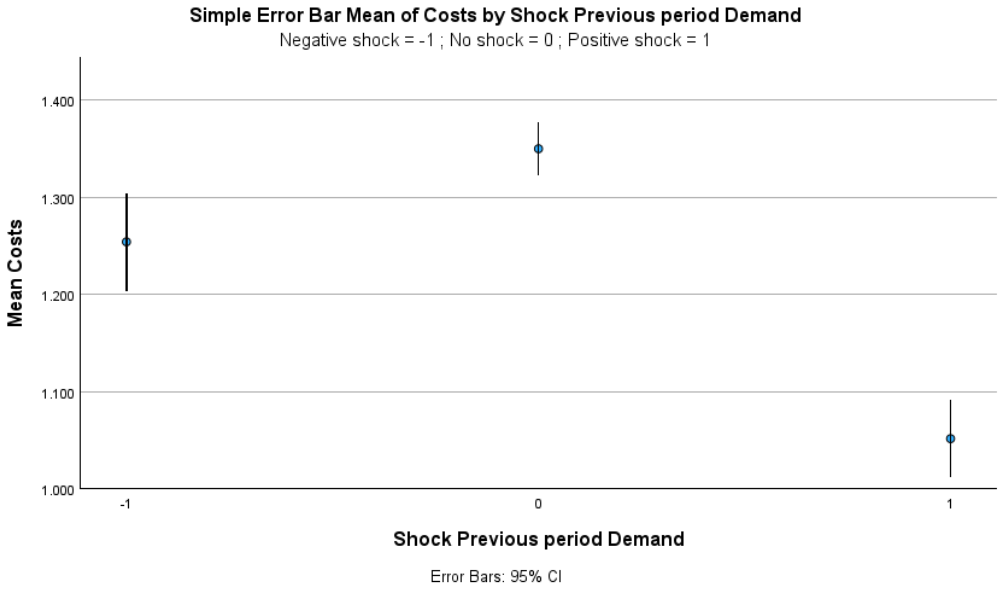

Figure 34: Simple Error Bar Shock Previous period Demand - Costs

6. A biometric analysis of the Newsvendor Problem

To make up for the lack of contributions to the study of behavioural variability in decision making, starting from the pertaining literature review, the importance of searching for an original method to study the problem was realised.

In this regard, the conducted literary revision ultimately led to the decision of analysing decision making from the innovative neurophysiological point of view, which is recently gaining the interest of many academicians.

Therefore, the present work aims to contribute to the described literary gap, thanks to an experiment designed and conducted (in its pilot form) in the IMPD Lab at the Politecnico di Torino.

In detail, the experiment, entirely conducted by the author of the present study, was based on the Newsvendor Problem with the addition of some tools for the detection of biometric variables, given that its purpose was to understand the mechanisms underlying decision making process, starting from the observation of the biometric characteristics of the decision maker.

Nevertheless, it was not possible to analyse the collected results because, in consideration of the voluntary nature of the participation to the experiment, a major limitation was found during the experimental phase: the lack of sufficient participants for the test. The rest of the present section of the paper therefore refers only to the design and implementation of the experiment, while for the analysis of the results, it is hoped that others will replicate the test, thus obtaining several records consistent with drawing conclusions from it.

6.1. Research questions and hypothesis

The experiment originated from the following four research questions:

1. *Are the literary conclusions about margin in decision making valid?*
2. *How does demand variability influence the behaviour of the decision maker?*
3. *Is there a connection between the decision maker's cognitive ability and his/her performance in decision making tasks and capacity to manage uncertainty?*
4. *How do margin, demand variability, and cognitive ability impact biometric measurement?*

Given the established consistency of the literary results concerning how the margin of the product influences the decision maker in the Newsvendor Problem [93], the first research question was conceived to assert the validity of the sample.

Instead, the other three questions represent the innovative part of the study, considering their current lack of investigation in the literary field.

Then, the following hypotheses were then designed:

1. *Order quantities tend to be above the optimal level in the low margin scenario and below optimal level in the high margin one.* [93]
2. *The higher the demand variability, the higher the deviation of the decision maker order quantity from the optimal level.* [125]
3. *The higher the decision maker score on the Cognitive Reflection Test (CRT), the lower his/her order quantity deviation from the optimal level.* [122]
4. *The lower the decision maker score in the overprecision test, (the lower his/her level of overprecision,) the lower his/her order quantity deviation from the optimal level.* [125]
5. *A "well-performing" decision maker (i.e. a participant who shows small deviations from optimal orders) shows higher anticipatory skin conductance levels preceding disadvantageous choices than advantageous ones (i.e., preceding negative outcomes than positive ones)* [126]
6. *The lower the decision maker inspiration-expiration rate, the lower his/her order quantity deviation from the optimal level (related to his/her higher HRV).* [127]
7. *The higher the decision maker heart rate variability (HRV), the lower his/her order quantity deviation from the optimal level.* [127]

In detail, the first hypothesis is referred to the first research question, the second to the second one, the third and the fourth to the third one, and the last three to the fourth one.

Except for the first one, all the formulated hypotheses refer to some aspects already investigated in the literature, but a deeper analysis is still required.

6.2. Sample

Participants were voluntary Logistics students from the Ingegneria Gestionale bachelor course of the Politecnico di Torino. Each participant required around two hours to be tested and analysed individually, per the criterion of anonymity.

Eight students initially showed interest in the experiment, but only five of these decided to book for the different available time slots of 1 hour and 45 minutes each. Until the day before the experiment, participants were not aware of the activities they would have performed, but only of the fact that they would have had the opportunity to implement managerial decisions, on which they would have later received some feedback. This choice was made to prevent people from specifically preparing for the execution of the Newsvendor Problem, given that the experiment aimed to test spontaneous behaviour. For this reason, once the student had arrived in the laboratory of the Politecnico di Torino, he/she was made aware of the structure of the test in detail, with particular reference to the Newsvendor experiment and the biometric measurements that would have been carried out in the meantime. After reading an explanatory document, complete freedom to choose not to consent to participate in the test was given, but all five students consented. In addition, the participant was also informed in advance of the possibility of deciding not to have his/her data taken into account in the research, even after the end of the test. In any case, the results of each participant have not been linked to the individual's identity[1].

6.3. Experimental methodology

The experiment required different months to be designed, needing a deep literary investigation to detail the research, with many different expedients aimed at reaching a comprehensive overview of how decision makers spontaneously behave in the presence of various elements: variability, diverse incidental emotional elements, distinct cognitive abilities, and many others.

6.3.1. Tools

Overall, the tools used during the test consisted of an Excel file in which participants inserted their answers; a PowerPoint presentation that gave timing information and displayed some questions to participants; the biometric equipment, consisting of a sensor for the cutaneous conductance, a

[1] This information was provided in the module reported in Appendix A.

respiratory band, an ECG sensor and a telemetric wireless sensor which communicated with the monitoring computer; and finally the Encephelan software for the biometric screening[2].

Participants were only asked to bring a paper sheet, a pen, and a calculator, to do some calculations in the way they preferred.

6.3.2. Overview of the experimental session

Once the participant signed the consensus, he/she was introduced to the biometric equipment, while it was placed on him/her. Then, the student was made to sit at the desk workstation, with one computer in front and another one next to him/her. Two different computers were in fact needed at the participant to execute the session: on the one they had in front, students had their Excel file with separate sheets for questionnaires and the Newsvendor Problem, while on the one next to them, a PowerPoint presentation. This setting was adopted to ensure a high level of automatization during the session, given that the PowerPoint slides reproducing questionnaires' questions and Newsvendor periods automatically changed without the need for the participants to do nothing. The only exception consisted in the slides of the transaction shown during breaks: in these cases, it was suggested to participants when to move on by clicking on the presentation. In addition, still, to automatize as much as possible the execution of the session, the PowerPoint presentation also reproduced sound signals which alerted the participants during each activity (i.e. a round of the Newsvendor, a question of a questionnaire, and so on), first when he/she had 10 seconds left for the specific activity and then when time was over. Only the third questionnaire was excluded from this setting since no time records were taken during it.

The author of the present work was instead sitting at another desk, giving her back to the students, monitoring their activity on two different screens: one checking online what students were doing on Excel and another monitoring their biometric parameters.

It is necessary to underline that both choices of not showing a countdown to participants (who were aware of the time spent only from the sounds reproduced by the presentation) and not directly watching them while performing, raised by the need not to intimidate the participant, who could have had fell unnecessary stress otherwise.

Thus, overall the experimental session was structured as reported in Figure 35.

[2] All these tools are further deepened in the subsequent sections.

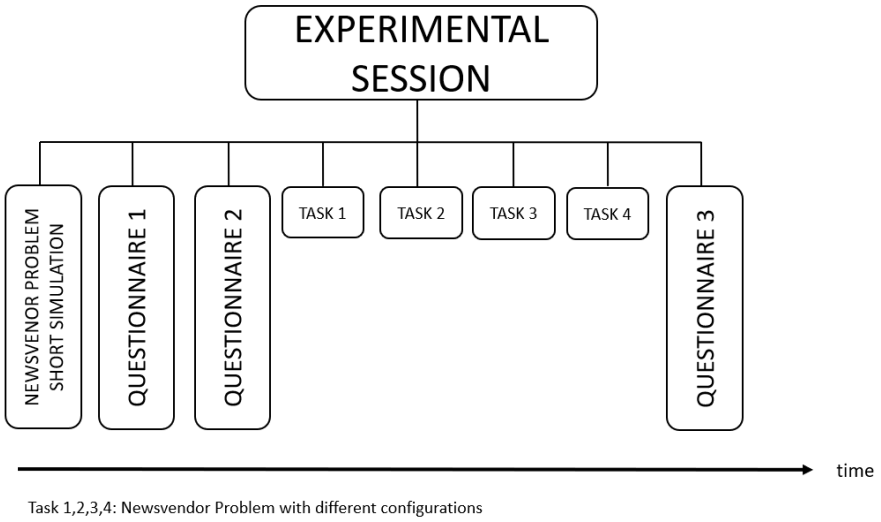

Figure 35: Structure of the experimental session

After a short simulation of few periods of a single Newsvendor task, participants were asked to compile two questionnaires: the Cognitive Reflection Test (CRT) and a test to measure their level of overprecision[3]. Then participants effectively started the Newsvendor Problem, and, after finishing that, they compiled the last questionnaire in which it was asked about their perceived performance and other information (i.e. academic performance and time gone since their last coffee and meal).

During the experimental session, each activity had a specific timing, with the exception of the last questionnaire, for which time monitoring (i.e. the subsequent calculation of the response time) was not required. In more detail, in Table 9 an overview of how time was managed is reported. Nevertheless, it has to be clarified that time management was not always very strict, because of different circumstances, such as small inconveniences occurred.

Table 9: Timing of the experimental session

ORDER OF THE ACITIVITIES	EXPERIMENTAL SESSIONS: 25/05/2023 AND 26/05/2023
1	Equipping the participant with biometric detection tools (15 min)
2	Newsvendor simulation on 4 rounds (4 min)
3	Questionnaire 1 (6 min)
4	Break (1 min)
5	Questionnaire 2 (6 min)
6	Break (1 min)
7	Task 1 (10 min)
8	Break (4 min)
9	Task 2 (10 min)
10	Break (4 min)
11	Task 3 (10 min)

[3] Both the CRT and the overprecision test are better explained in the "Data collected" chapter.

12	Break (4 min)
13	Task 4 (10 min)
14	Questionnaire 3 (2 min)
15	Removing the biometric detection tools from the participant (15 min)

6.3.3. Data collected

In accordance with the availability of the University's tools for the biometric measurements, and based on the literature current investigations, the data that it was decided to measure are reported in Table 10.

Table 10: Data collected during the experiment

PARTICIPANT CHARACTERISTICS	BEHAVIOURAL VARIABLES	BIOMETRIC PARAMETERS
Cognitive Reflection Test (CRT)	Demand forecast	Electrodermal activity (EDA)
Overprecision	Purchase quantity	Respiratory rate
Academic performance	Response time	Heart rate variability (HRV)
Perceived performance	/	/
Time since last coffee and meal	/	/
Respiratory/Heart disease		

6.3.3.1. Cognitive Reflection Test

The Cognitive Reflection Test, already discussed in the previous chapter, usually encounters the so-called familiarity issue, since people are often familiar with it, compromising the reliability of the results, considering that it is specifically adopted to test spontaneous behaviour. For this reason, instead of the classic three-questions formulation, the longer six-questions version of the test formulated by Primi et al. was used for the experiment, as presented below (the first three questions belong to the original CRT formulated by Frederick and the other three are the additional items included by Primi et al.) [128]:

1. *A bat and a ball cost €1.10 in total. The bat costs €1.00 more than the ball. How much does the ball cost?* [correct answer = 5 cents; heuristic answer = 10 cents]
2. *If it takes 5 minutes for five machines to make five widgets, how long would it take for 100 machines to make 100 widgets?* [correct answer = 5 minutes; heuristic answer = 100 minutes]

3. *In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?* [correct answer = 47 days; heuristic answer = 24 days]
4. *If three elves can wrap three toys in hour, how many elves are needed to wrap six toys in 2 hours?* [correct answer = 3 elves; heuristic answer = 6 elves]
5. *Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are there in the class?* [correct answer= 29 students; heuristic answer= 30 students]
6. *In an athletics team, tall members are three times more likely to win a medal than short members. This year the team has won 60 medals so far. How many of these have been won by short athletes?* [correct answer = 15 medals; heuristic answer = 20 medals]

6.3.3.2. Overprecision test

Participants' level of overprecision was also considered of interest for the study, given the already discussed positive correlation between the bias and the performance level in the Newsvendor Problem that emerged from the literature review [119]. Many possible alternatives to test the bias were evaluated, in the end deciding for the one proposed by Ren, D. Croson and R. Croson in 2017, which consist in two questions asking participants to estimate the 25th and 75th percentile of a uniform distribution [125]. However, the literary approach was slightly modified by adding two other questions, in which participants had to provide the same estimates but of a not-uniform, discrete distribution, represented by a randomly generated numeric string. This innovative content was added in order to test also how participants deal with a probably less-usual but more realistic scenario. Following the four questions asked to participants are presented, with the last two being the additional ones:

1. *Consider a uniform distribution with a lower bound of 1 and an upper bound of 300 (i.e. a distribution in which each number between 1 and 300 has the same probability to be extracted). Which is your estimate of the 25th percentile of this distribution? (The 25th percentile is a number such that 25th of the time, a random draw from this distribution is less than this number)*
2. *Consider a uniform distribution with a lower bound of 1 and an upper bound of 300 (i.e. a distribution in which each number between 1 and 300 has the same probability to be extracted). Which is your estimate of the 75th percentile of this distribution? (The 75th*

percentile is a number such that 75th of the time, a random draw from this distribution is less than this number)

3. *Consider the following numeric string: 25, 30, 36, 47, 51, 59, 65, 68, 75, 81, 93, 99. Which is your estimate of the 25th percentile of this string?*
4. *Consider the following numeric string: 25, 30, 36, 47, 51, 59, 65, 68, 75, 81, 93, 99. Which is your estimate of the 75th percentile of this string?*

6.3.3.3. Third questionnaire

As the last activity of the session, participants were instead asked to specify their academic performance, their perceived performance and the amount of time gone since their last coffee and meal.

It was hence requested that students to provide their academic average and the number of sustained exams, in order to investigate a possible positive correlation between university performance and Newsvendor decisions.

The second data was instead explored by asking: "*In which quartile do you place yourself with respect to other participants?*". This question was asked without referring to the total amount of students because each one was examined individually, not giving further information about the others. For clarity, four different quartiles were proposed, with attached their definition, namely: first quartile – better than 75% of the others, second quartile – better than 50% of the others, third quartile – worse than the 50% of the others, and fourth quartile – worse than the 75% of the others.

Finally, the information about the last coffee and meal was asked of participants to guarantee a sufficiently critical lecture on the biometric measurements.

6.3.3.4. Behavioural variables

Moving on to the central column of Table 10, the demand forecast and purchase quantity points refer specifically to the Newsvendor Problem activity, in which participants were required to provide both demand forecast and purchase values. This setting was decided to have a complete vision of what is affected by biasing dynamics, both at the level of purchase and of the forecast.

Response time was instead collected both during the execution of the first two questionnaires and of the Newsvendor Problem. Participants' Excel files were automatized and able to register the exact time the answers were given in hidden columns that participants could not see during the test. This

expedient was realized so that participants did not feel the pressure of the passage of time. On the other hand, the starting time of each activity was collected by the author of this paper in order to be able to calculate the response time as the difference between the two moments.

6.3.3.5. Biometric parameters

Three biometric parameters were registered during the entire session.

Electrodermal activity (EDA) was measured through a skin conductance sensor, consisting of two adjustable rings of elastic material inserted in two fingers of the participant's left hand, with two electrodes attached. In addition, also a "ground wire" was placed on the participant to have a neutral benchmark for analysing skin parameters through an adhesive electrode placed on their left forearm.

Respiratory rate was traced through a chest/abdominal breathing sensor that consists of an adjustable band placed around the participant's chest (without the need for the participant to remove their upper garments). The band is adjustable, must be sufficiently tight to the chest to allow the correct tracking of the respiratory rate, and presents a rigid sensor on the front of the band.

Finally, an electrocardiogram was recorded with a sensor consisting of three electrode pads placed on both the participant's wrists and one on the left forearm to measure their heart rate variability (HRV).

The electrodes and the rigid sensor of the abdominal band were connected to a wireless telemetric recorder, which served as a a link between the participant and the Encephelan software. While participants were executing their session, all the data were correctly reproduced on the monitor.

Moreover, about the registered parameters, it was also asked participants, before starting the test, if they had any respiratory or heart disease to ensure correct analysis of the collected biometric parameters.

6.3.4. The Newsvendor Problem execution

Each task of the Newsvendor Problem was structured as shown in Figure 36.

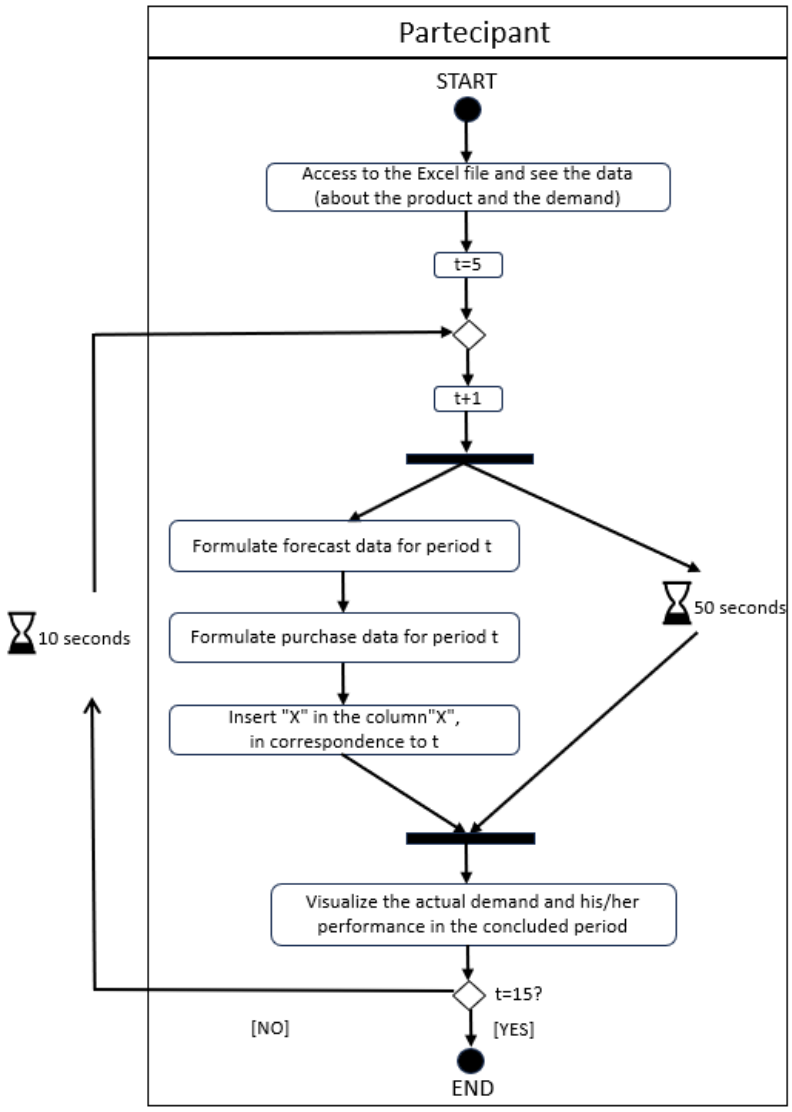

Figure 36: Structure of the Newsvendor Problem tasks

The problem was divided into 4 tasks of 15 rounds each, with known demand data for the first 5 periods. This was made in order to give a starting overview of the actual demand and see how participants related to it. Participants had to provide demand forecast and purchase values for each of the 10 rounds. Additionally, when participants formulated their official period values, they were asked to insert an "X" in a dedicate column to make their performance values and demand data appear, thanks to the automation provided to the Excel file. This system was invented, again, in order to guarantee sufficient automation during the test, letting participants discover by themselves the values of performance and actual demand. Indeed, during the Newsvendor Problem execution (differently from the first two questionnaires), the PowerPoint presentation was only needed to hear the time signals and to underline the current round, since everything the participant needed was displayed on his/her Excel file (as in the last questionnaire).

After entering the "X", participants were instructed not to modify their answers for the concluded round, moving on to the subsequent period in accordance with time specifications. In Figure 37 the configuration of each task is reported, with participants being able to modify only the green columns, inserting values and using Excel formulas if they wanted to. In detail, the yellow column was in charge of automatically showing demand data for each period, and the red ones the performance values, automatically after the "X" input.

Period	Forecast	Purchase	X	Purchase costs	Demand	Sales volumes	Overstock volumes	Profit
1					192			
2					41			
3					81			
4					103			
5					191			
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								

Figure 37: Participants' task Excel configuration

6.3.4.1. Experimental variables

Two experimental variables were manipulated during the Newsvendor Problem execution, namely product margin and demand variability [93].

A third value, that is framing, was also considered in the phase of experimental design. Given that only some initial investigations are currently present in literature about the topic (e.g. Schultz et al., 2018 [120]), how information framing influences participants is an interesting field to explore. However, given the higher number of participants, the longer time and the grater complexity that it would have had required, it was ultimately decided not to take it into account. This was made by not differentiating the tasks based on the way in which information was framed to participants.

For each task, the following information was given to participants: Sale price, Overstock sale price and Purchase cost [93]. In detail, the first and the third task, were realized with a high margin product, while the second and the fourth ones with a low margin product, respectively characterized by the following unitary values: the Sale price was 180€ in the high margin scenario and 120€ in the low margin one; the Overstock sale price was 80€ for the high margin product and 20€ for the low margin one. The Purchase cost was instead €100 / unit in both cases.

The fact that these specific information was given to participants makes evident that the proposed framing was positive (by many literary studies), but not changing it between tasks made its impact analysis unavailable.

The four tasks also differed in demand variability, even if participants were unaware of this. Hence, while the mean demand was 100 in all four cases (in accordance with the experiment conducted by Bolton and Katok in 2008 [112]), the first two tasks were characterized by high demand variability, and the last two by low demand variability. This was produced by modifying the values of demand standard deviation, which was equal to 50 for the high variability tasks and to 20 for the low variability ones. These values were founded in literature and slightly modified to create larger differences among the two scenarios. In 2014 Sachs and Minner realized in fact a Newsvendor study with demand coefficients of variations[4] equal to 0.5 and 0.3 [129]. These values were therefore adopted, but modified in 0.5 and 0.2.

The demand values for all 20 periods were thus elaborated in accordance with the criteria mentioned, by imposing these on a random number generator created on Excel, which produced different values per task and participant.

For major clarity, an overview of what was stated above about the differences among the tasks is reported in Table 11. The represented scenarios were chosen in order to evaluate how decision makers relate to different product margin and demand variability conditions, understanding how these different factors influence decisions.

Table 11: Product and Demand values per task

TASK	PRODUCT				DEMAND		
	MARGIN	PURCHASE COST	SALE PRICE	OVERSTOCK SALE PRICE	VARIABILITY	MEAN	STANDARD DEVIATION
1	HIGH	100€	180€	80€	HIGH	100	50
2	LOW	100€	120€	20€	HIGH	100	50
3	HIGH	100€	180€	80€	LOW	100	20
4	LOW	100€	120€	20€	LOW	100	20

[4] The ratio between the standard deviation and the mean.

6.4. Feedback

After some days from the test, participants received the promised feedback about the decisions taken during the experiment. A detailed document was prepared for each student, maintaining the anonymous setting[5].

The document consisted of three separate sections, respectively dedicated to the CRT, the Newsvendor Problem execution, and some conclusions about the general performance.

All the questions of the CRT were thus presented, each with annexed the literary correct and heuristic answers, and the answer inserted by the candidate [128].

Each participant was then classified as a "Slow thinker" or a "Fast thinker", depending on the average score obtained on the test, according to the rule that if three or more of his/her answers were correct, he/she was considered a "Slow thinker", or a "Fast thinker" conversely.

In many cases, students answered the CRT with some values neither correct nor heuristic, but these answers were considered incorrect. Then, students also obtained some specific insights about the managerial decisions taken during the Newsvendor Problem.

In particular, a summary of their purchase and forecast decisions was associated with the corresponding levels of distance from the actual demand, the optimum order and the optimum profit, separately in different tables. A focus was specifically dedicated to the comprehension of how the candidate performed in different experimental conditions.

Finally, in the last section, some short conclusions about the general behaviour of the participant were reported.

[5] An exemplary feedback document is reported in Appendix B.

Conclusions

Overall, the analytical analysis conducted for the present study confirmed the widely established literary finding that individuals are prone to be influenced by various circumstances when deciding, systematically deviating from optimal behaviours.

In detail, among the possible causes of behavioural variability, the specific focus conducted on the Newsvendor Problem highlighted how experimental conditions, individual cognitive abilities and (briefly) previous period conditions differently impact decision makers when they forecast demand, make purchases, and incur costs as a result.

To sum up, some conclusions about the overall analysis can be drawn by looking at the hypothesis that the study aimed to test.

The concept that individuals vary their purchase attitude based on product margin and information framing, entailed in the first two hypothesis, was accepted. Indeed,, the literary results about the Pull-to-centre effect and Prospect theory were confirmed.

Demand forecast appeared instead to be influenced by margin and not by framing, exactly contrary to what was stated in the third hypothesis, being this result however reputed acceptable because of the exploratory nature of the topic and the low number of contributions currently investigating it.

Another variable that resulted fundamental to be taken into account is the specific experimental session, which indeed proved the occurrence of a learning effect in the examined sample, hence demonstrating the fourth hypothesis to be right. Moreover, the data observation across sections, also evidenced that, despite an overall risk-aversion, individuals acquired confidence over time, both forecasting and purchasing higher values in the second session than in the first one.

Overall, the scenario characterized by low product margin and positive information framing was the one in which participants performed better in terms of deviations from best practices (both of purchases and costs).

For what concerns instead how individuals are affected by their cognitive abilities in the decision making process, the submission of the Cognitive Reflection Test resulted in a valuable expedient to understand newsvendors' performance. However, the test score appeared unable to significantly influence the variability of the orders placed by the newsvendors, unless it was associated with other forms of cognitive ability, thus proving the fifth hypothesis to be correct, but with reserve.

Nevertheless, the CRT score emerged as considerably negatively related to the average costs sustained by participants, proving its relevant role as an individual performance indicator.

Given the university nature of the sample, another variable that was widely evaluated as cognitive ability is the participants academic average, which confirmed its valuable role in determining deviations from optimal costs, as stated in the sixth hypothesis. However, overall, the supposition was only partially true, since participants' academic path did not appear significantly relevant in determining the discrepancy between the optimal orders and the participants' ones. Moreover, specific evidence that emerged from the evaluation of the variable in its binary form (using twenty-five as a threshold mean) is that, despite the lower costs sustained, on average the respondents with a higher academic mean were characterized by a greater answering variability. In detail, this was considered highly relevant for the analysis since it strongly emphasizes how multifaceted behavioural variability is and, thus how it has to be carefully studied without giving anything for granted.

The study of cognitive abilities also entailed a last variable as possible source of variability: participants' perceived performance, which indicates the quartile in which participants placed their task performance. Overall, the study of the variable brought to the rejection of the seventh hypothesis, given that it did not result in influencing either the deviations from the optimal purchases, or the ones from the optimal costs. However, despite the lack of a statistically relevant influence, it was assessed that the pattern identified in the hypothesis actually existed concerning costs. Indeed, a sort of cognitive ability was found in the "medium" answering behaviour: participants who placed their own performance in the average quartiles, were the ones who sustained the lowest costs.

However, the perceived performance was evaluated also by looking at the differences between participants' perceived and actual performance. In turn, this additional perspective resulted more relevant in impacting costs as an indicator of actual good performance (i.e. when its value is higher than zero), because of its negative relationship with costs. The analysis also highlighted the wider variability characterizing the lower values of the measure, underlying (at least in this specific case) the connection between the scarcity of cognitive ability and the behavioural variability.

Then, in order to link all the previously made observations, it was specifically deepened on how the different variables impact the standard deviation values of both purchase quantities and costs and how these, in turn, are reciprocally related. In this regard, the results of the analysis proved the eighth hypothesis to be right and the detrimental role of behavioural variability: the higher the dispersion

of the demand forecast and purchase quantities values, the higher the average costs sustained by participants, thus the worst their performance.

A different analysis was instead conducted to investigate how the realization of a shock in the previous period demand influenced participants in incurring in costs. In doing so, the ninth hypothesis was confirmed to be acceptable, given that a relevant impact was proved. Moreover, the analysis also showed two other interesting results: first, despite the overall negative correlation between the variables, the highest costs characterized the no-shock scenario, probably suggesting that people are more prone to modify their conduct when they are subjected to a strong modification in the external inputs; on the other hand, the fact that the highest variability levels were registered in case of shock points out how the variety of individual behaviours also determines how people react, and again the fact that behavioural variability does not leave room for obvious considerations.

Limits and future steps

The very last part of the study was aimed at proposing an original experimental setting based on the observation of participants' biometric characteristics, which has been implemented in its pilot form for this work. However, due to the voluntary nature of the study, not enough participants were found on time to draw some statistically relevant conclusions. Thus, if others will want to contribute to the study of the topic, the data collection will be continued in the future.

Appendix A

POLITECNICO DI TORINO

SESSIONE SPERIMENTALE PER L'INVESTIGAZIONE DELLE DECISIONI NEL NEWSVENDOR PROBLEM:

REPORT SULLE DECISIONI DEL CANDIDATO 2

Nota introduttiva:

*La dicitura "Candidato *numero*" riportata nel presente documento fa riferimento esclusivamente all'ordine in cui i partecipanti della sessione sperimentale sono stati esaminati, non rappresentando quindi una classifica realizzata in base al punteggio ottenuto.*

Il presente documento è inoltre stato realizzato nel rispetto del criterio di anonimità, senza possibilità di ricondurre le informazioni riportate allo specifico partecipante.

Nel presente documento sono riportati i feedback sulle decisioni prese nel corso della sessione sperimentale svolta in data 25/05/2023 presso gli uffici del Politecnico di Torino. Il documento consta di tre parti, rispettivamente dedicate al Questionario 1, al Newsvendor Problem e alle conclusioni sul comportamento del partecipante.

1. QUESTIONARIO 1

Il primo questionario svolto nel corso della sessione sperimentale consisteva nel noto Cognitive Reflection Test (CRT).
Descritto per la prima volta dallo psicologo Shane Frederick nel 2005, il CRT è un test progettato per valutare la tendenza di una persona ad ignorare una risposta errata istintiva e ad impegnarsi in una più profonda riflessione per trovare la risposta corretta ad alcune domande.

Nel corso della sessione sperimentale è stata utilizzata la formulazione più lunga del CRT, realizzata da Primi et. al. nel 2016, caratterizzata dalle sei (invece che tre) domande riportate di seguito.
Per ciascuna domanda del questionario esiste una risposta corretta ("Slow", i.e. ragionata e corretta) ed una risposta euristica ("Fast", i.e. più immediata ma scorretta), indicate alla fine di ogni domanda, insieme alle risposte date dal candidato.
I colori delle risposte del candidato indicano rispettivamente: il verde le risposte corrette, il rosso quelle euristiche e l'azzurro quelle scorrette ma non appartenenti a nessuna delle due categorie sopra citate.
Al termine del questionario è infine indicata la valutazione complessiva del candidato.

1. *Una mazza e una palla costano complessivamente 1,10 euro. La mazza costa 1,00 euro in più della palla. Quanto costa la palla?*

Risposta corretta	Risposta euristica	Riposta partecipante
€0.05	€0.10	€0.10

2. *Se cinque macchine impiegano 5 minuti per produrre cinque oggetti, quanto tempo impiegheranno 100 macchine per produrre 100 oggetti?*

Risposta corretta	Risposta euristica	Riposta partecipante
5 min	100 min	50 min

3. *In un lago c'è un frammento di ninfee. Ogni giorno il frammento raddoppia di dimensioni. Se il frammento impiega 48 giorni per coprire l'intero lago, quanto tempo impiegherà il frammento per coprire metà del lago?*

Risposta corretta	Risposta euristica	Riposta partecipante
47 giorni	24 giorni	47 giorni

4. *Se tre elfi possono incartare tre giocattoli in un'ora, quanti elfi sono necessari per incartare sei giocattoli in 2 ore?*

Risposta corretta	Risposta euristica	Riposta partecipante
3 elfi	6 elfi	3 elfi

5. *Jerry ha ricevuto sia il 15° voto più alto che il 15° più basso della classe. Quanti studenti ci sono nella classe?*

Risposta corretta	Risposta euristica	Riposta partecipante

29 studenti	30 studenti	29 studenti

6. *In una squadra di atletica, i membri alti hanno tre volte più probabilità di vincere una medaglia rispetto ai membri bassi. Quest'anno la squadra ha vinto finora 60 medaglie. Quante di queste sono state vinte da atleti bassi?*

Risposta corretta	Risposta euristica	Riposta partecipante
15 medaglie	20 medaglie	20 medaglie

Valutazione complessiva partecipante:

Tabella 1: Valutazione candidato in base al CRT

Totale risposta corrette	Totale risposte euristiche	Totale riposte "altre"
3	2	1

Alla luce delle risposte date, il candidato è stato identificato come: **SLOW THINKER**

Lo "Slow Thinker", come evidenzia lo psicologo Daniel Kahneman nel libro "Thinking, Fast and Slow" del 2011, è un individuo che tende ad affidarsi principalmente alla propria sfera logica e riflessiva, risolvendo quindi problemi in maniera ragionata e non impulsiva, giungendo a delle scelte ben ponderate e spesso corrette.

2. NEWSVENDOR PROBLEM

Nel seguente paragrafo sono invece riportate le performance ottenute dal candidato nel corso dell'esperimento del Newsvendor.
Il problema del Newsvendor (o del Newsboy) è un problema di gestione delle scorte che inizia ad essere studiato intorno agli anni '50, divenendo successivamente di interesse per l'investigazione di alcuni aspetti comportamentali legati alle decisioni.
In ogni sua formulazione, il problema prevede che un decisore (i.e. il newsvendor) debba acquistare ad un prezzo c una quantità q di un dato prodotto che rivenderà ad un prezzo p nel corso del periodo di vendita, senza però sapere l'effettiva domanda di questo.
Il problema prevede un valore ottimo per gli acquisti effettuati dal partecipante calcolato tramite la Formula 1 (Brandimarte e Zotteri, 2007), basata sulla media e sulla varianza della domanda di mercato, e sul margine del prodotto.

$$F(Q*) = \frac{m}{m+c}$$

Formula 1: Brandimarte e Zotteri, 2007

Questa formula permette di individuare la quantità ottima Q^* in funzione di c e p. Poiché lo scopo dell'esperimento è quello di investigare gli aspetti legati alle decisioni umane, durante

lo svolgimento dell'esperimento non era tuttavia necessario che il partecipante applicasse questa formula.

2.1 L'esperimento

Nella formulazione proposta durante la sessione sperimentale, il partecipante è stato sottoposto allo svolgimento di 4 tasks, da 15 periodi ciascuno, con domanda nota per i primi 5 periodi. Per ciascun periodo, al partecipante veniva quindi richiesto di specificare i propri valori di previsione della domanda e acquisti, entro 50 secondi.
Inoltre, tutti i tasks erano caratterizzati da framing positivo, in quanto al partecipante venivano comunicati esclusivamente i valori di: prezzo di vendita, prezzo di svendita e costo di acquisto. Tuttavia i tasks differivano tra loro per i diversi valori delle variabili sperimentali manipolate nel corso dello studio, ovvero:

- il margine del prodotto – alto per un prodotto con un'evidente differenza tra prezzo di vendita e costo di acquisto; basso per un prodotto con una differenza non molto elevata;
- la variabilità della domanda – alta se caratterizzata da una deviazione standard elevata; bassa se caratterizzata da una deviazione standard non molto elevata

Di seguito è riportato uno schema di come le variabili sono state gestite nei diversi tasks:

Tabella 2: Variabili sperimentali Newsvendor Problem

TASK	PRODOTTO				DOMANDA		
	MARGINE	COSTO DI ACQUISTO	PREZZO DI VENDITA	PREZZO DI SVENDITA	**VARIABILITA'**	MEDIA	DEVIAZIONE STANDARD
1	ALTO	100 €	180 €	80 €	ALTA	100	50
2	BASSO	100 €	120 €	20 €	ALTA	100	50
3	ALTO	100 €	180 €	80 €	BASSA	100	20
4	BASSO	100 €	120 €	20 €	BASSA	100	20

Alla luce di ciò, le seguenti definizioni riportate nelle pagine successive sono da intendersi come segue:

- "margine alto" : media dei valori di primo e terzo task
- "margine basso" : media dei valori di secondo e quarto task
- "variabilità alta" : media dei valori di primo e secondo task
- "variabilità bassa" : media dei valori di terzo e quarto task

Di seguito sono dunque riportati i risultati conseguiti dal candidato, in termini di: profitto conseguito; scostamento dalla domanda, dall'ordine ottimo e dal profitto ottimo; tasso di miglioramento tra un task e il successivo (il segno che caratterizza il tasso di miglioramento indica un peggioramento nel caso di segno negativo o un effettivo miglioramento nel caso di mancanza di segno).

In aggiunta alle performance del partecipante, sono inoltre presentati per ciascun periodo i valori di domanda, ordine ottimo e profitto ottimo.

2.2 Acquisti

Nella Tabella 3 sono riportati i valori di Previsione della domanda e Acquisti inseriti dal partecipante e l'effettiva realizzazione della domanda, per i 10 periodi di ciascuno dei 4 tasks affrontati dal partecipante nel corso dell'esperimento del Newsvendor.
In aggiunta, per ciascun task, sono stati calcolati i valori medi di Acquisti, Previsione e Domanda sui 10 periodi.

A partire da tali valori medi, sono inoltre stati riportati gli scostamenti per task dalla Domanda, sia a livello di Acquisti che di Previsione.

Tabella 3: Previsione domanda, Acquisti effettuati e Domanda

ROUND	TASK											
	1			2			3			4		
	PREVISIONE	ACQUISTI	DOMANDA	PREVISIONE	ACQUISTI	DOMANDA	PREVISIONE	ACQUISTI	DOMANDA	PREVISIONE	ACQUISTI	DOMANDA
6	86	90	62	154	150	61	100	100	79	22	20	92
7	105	110	10	91	50	101	76	50	98	131	100	106
8	110	100	108	150	130	152	61	50	119	133	110	111
9	130	120	159	18	20	35	95	70	103	63	50	106
10	86	90	120	70	80	138	74	65	98	77	67	99
11	79	80	166	166	150	26	79	75	133	88	80	84
12	18	20	121	110	30	138	27	25	129	101	95	90
13	130	110	133	166	100	110	103	80	65	108	103	98
14	125	100	65	42	30	161	117	100	103	118	110	97
15	100	105	41	27	20	83	125	100	119	78	60	68
MEDIA PER TASK	**96,9**	**92,5**	**98,5**	**99,4**	**76**	**100,5**	**85,7**	**71,5**	**104,6**	**91,9**	**79,5**	**95,1**
SCOSTAMENTO % DA DOMANDA PER TASK	**-1,62%**	**-6,09%**		**-1,09%**	**-24,38%**		**-18,07%**	**-31,64%**		**-3,36%**	**-16,40%**	

Analizzando i dati si nota che mediamente in tutti i task il candidato si è distanziato meno dall'effettiva domanda tramite i valori di previsione, rispetto a quelli di acquisto, probabilmente a causa di una strategia volta a ridurre al minimo i possibili volumi di svendita. Ad ogni modo, mediamente il valore di domanda non è in nessun task stato oltrepassato né dalla previsione, né dagli acquisti.
Si riscontra inoltre che il partecipante non sembra essersi lasciato influenzare negativamente dall'alta variabilità della domanda, specialmente nel primo scenario in cui ha complessivamente gestito meglio che negli altri casi sia la previsione che gli acquisti.

Nella Tabella 4 sono invece riportati gli Acquisti effettuati dal partecipante e l'Ordine ottimo (i.e. quello che avrebbe portato alla massimizzazione del profitto) per i 10 periodi di ciascuna dei 4 tasks affrontati dal partecipante nel corso dell'esperimento del Newsvendor.
In aggiunta, per ciascun task, sono stati calcolati i valori medi di Acquisti e Ordine ottimo sui 10 periodi.
A partire da tali valori medi, sono inoltre stati riportati i valori di Acquisti e Ordine ottimo medi, distinti per margine di prodotto (alto nei tasks dispari e basso nei pari).
Infine, sia per ciascun task che sulla base del margine del prodotto, sono stati calcolati gli scostamenti percentuali degli Acquisti medi dall'Ordine ottimo medio. Uno scostamento negativo indica quindi che l'Acquisto medio è stato inferiore all'Ordine ottimo e viceversa uno scostamento positivo che è stata ordinata una quantità di merce superiore all'Ordine ottimo.

Tabella 4: Acquisti effettuati e Ordini ottimi

ROUND	TASK							
	1		2		3		4	
	ACQUISTI	OTTIMO	ACQUISTI	OTTIMO	ACQUISTI	OTTIMO	ACQUISTI	OTTIMO
6	90	138	150	47	100	101	20	84
7	110	131	50	52	50	102	100	86
8	100	131	130	59	50	108	110	88
9	120	141	20	50	70	108	50	90
10	90	142	80	55	65	108	67	90
11	80	149	150	47	75	114	80	88
12	20	149	30	52	25	118	95	87
13	110	150	100	55	80	116	103	87
14	100	146	30	58	100	116	110	87
15	105	143	20	59	100	117	60	84
MEDIA PER TASK	**92,5**	**142**	**76**	**53,4**	**71,5**	**110,8**	**79,5**	**87,1**
ACQUISTO MEDIO vs OTTIMO MEDIO MARGINE ALTO	**82 vs 126,4**							
ACQUISTO MEDIO vs OTTIMO MEDIO MARGINE BASSO	**77,75 vs 70,25**							
SCOSTAMENTO % DA ORDINE OTTIMO PER TASK	**-34,86%**		**42,32%**		**-35,47%**		**-8,73%**	
SCOSTAMENTO % DA ORDINE OTTIMO MARGINE ALTO	**-35,13%**							
SCOSTAMENTO % DA ORDINE OTTIMO MARGINE BASSO	**10,68%**							

Quanto riportato in Tabella 4 conferma i risultati letterari relativi al Pull-to-center-effect, la tendenza degli ordini ad essere al di sopra della quantità media ottima nel caso di prodotto a basso margine e al di sotto della stessa nel caso di margine alto (Schweitzer and Cachon, 2000).
Inoltre, gli scostamenti dall'ordine ottimo per task mostrano che, sebbene (come visto in Tabella 3) il candidato si sia distanziato meno dalla domanda con i propri acquisti nel primo task, lo stesso non vale per la distanza dall'ottimo, che risulta infatti complessivamente minore nell'ultimo task, scenario caratterizzato da basso margine di prodotto e bassa variabilità della domanda.

2.3 Profitto

Infine, la Tabella 5 e la Tabella 6 si focalizzano sul Profitto conseguito dal partecipante e sul Profitto ottimo (i.e. il massimo profitto che sarebbe scaturito dall'ordine ottimo) per i 10 periodi di ciascuno dei 4 tasks affrontati dal partecipante nel corso dell'esperimento del Newsvendor.
Per ciascun task, sono stati calcolati i valori medi di Profitto e Profitto ottimo sui 10 periodi. A partire da tali valori medi, sono inoltre stati riportati i valori di Profitto e Profitto ottimo medi, distinti per margine di prodotto e variabilità della domanda (alta nelle prime due tasks e bassa nelle ultime due). Sia per ciascun task che sulla base del margine del prodotto e della variabilità della domanda, sono poi stati calcolati gli scostamenti percentuali del Profitto medio dal Profitto ottimo medio. Diversamente da quanto osservato nella Tabella 4, in questo caso lo scostamento risulta sempre positivo, indicando però la percentuale assoluta di distanziamento dal Profitto ottimo.
Infine, sulla base degli scostamenti medi dal Profitto ottimo, viene presentato il Tasso di miglioramento del partecipante tra un task e il successivo. Tale dato è da intendersi come un

miglioramento della performance nel caso di positività del valore e come un peggioramento nel caso di negatività. Rispettivamente, i due casi indicheranno infatti che si è riscontrato un incremento nello scostamento medio tra un task e l'altro o un decremento.

Tabella 5: Profitto conseguito e Profitto ottimo: focus su task

ROUND	TASK							
	1		2		3		4	
	PROFITTO	OTTIMO	PROFITTO	OTTIMO	PROFITTO	OTTIMO	PROFITTO	OTTIMO
6	4400	3440	-5900	2620	5900	5880	400	2640
7	-1200	-1620	1000	6920	4000	7760	2000	4120
8	8000	8180	2600	12340	4000	10620	2200	4520
9	9600	14520	400	-500	5600	8140	1000	3720
10	7200	9160	1600	11060	5200	7640	1340	2880
11	6400	14980	-9400	-1160	6000	12540	1600	1360
12	1600	9120	600	11360	2000	11420	1400	2100
13	8800	10300	2000	7700	4900	4180	1560	3060
14	4500	3580	600	13520	8000	7980	900	2940
15	2000	1240	400	4060	8000	9720	1200	80
MEDIA PER TASK	**5130**	**7290**	**-610**	**6792**	**5360**	**8588**	**1360**	**2742**
SCOSTAMENTO % DA PROFITTO OTTIMO PER TASK	**29,63%**		**108,98%**		**37,59%**		**50,40%**	
TASSO DI MIGLIORAMENTO TRA TASK			**-79,35%**		**71,39%**		**-12,81%**	

Dal punto di vista del profitto conseguito il margine del prodotto sembra invece aver avuto una maggiore influenza nella realizzazione di profitti alti e scostamenti bassi dall'ottimo, rispetto alla variabilità della domanda. In particolare, la favorevole condizione di alto margine di prodotto caratterizzante il primo e il terzo task, ha condotto il partecipante ad ottenere i profitti complessivamente maggiori e gli scostamenti dall'ottimo minori, come si evince anche dall'osservazione dei tassi di miglioramento tra task.

Tabella 6: Profitto conseguito e Profitto ottimo: focus su caratteristiche sperimentali

PROFITTO	PARTECIPANTE	OTTIMO	SCOSTAMENTO %
MEDIO MARGINE ALTO	5245	7939	**33,93%**
MEDIO MARGINE BASSO	375	4767	**92,13%**
MEDIO VARIABILITA' ALTA	2260	7041	**67,90%**
MEDIO VARIABILITA' BASSA	3360	5665	**40,69%**

Quanto discusso fino ad ora, si riconferma in conclusione nella Tabella 6: alto margine di prodotto e bassa variabilità della domanda sono state le condizioni sperimentali in cui il partecipante è riuscito a performare meglio, distanziandosi meno dall'ottimo.

3. Conclusioni

Nel complesso il partecipante ha dimostrato la capacità di gestire diverse tipologie di scenari e di non lasciarsi influenzare facilmente da condizioni avverse, quali l'alta variabilità della domanda. Tuttavia, le performance migliori sono state registrate nelle favorevoli condizioni di alto margine di prodotto e bassa variabilità della domanda, che hanno complessivamente condotto il candidato a distanziarsi meno rispettivamente dal profitto ottimo e dall'ordine ottimo.

Appendix B

Politecnico di Torino

FOGLIO INFORMATIVO

PROTOCOLLO DI STUDIO DI PARAMETRI BIOMETRICI DURANTE LA RISOLUZIONE DI PROBLEMI DI GESTIONE DELLE SCORTE

Gentile interessato/a,
intendiamo proporle di partecipare ad una ricerca e, al fine di informarla circa lo scopo e le caratteristiche della ricerca stessa affinché lei possa decidere in modo consapevole e libero se partecipare, la invitiamo a leggere attentamente quanto riportato di seguito. I ricercatori coinvolti in questo progetto sono a disposizione per rispondere alle sue eventuali domande.

Responsabile scientifico dello studio

---

1. <u>Qual è lo scopo di questo studio?</u>

Lo studio intende indagare l'evoluzione di alcuni parametri biometrici del partecipante, mentre questo svolge il problema del Newsvendor e mentre risponde ad alcuni questionari.

Tale studio adotta l'approccio della fisiologia e della psicologia per la registrazione e raccolta di dati fondamentali riguardanti i processi neurofisiologici e psicologici del partecipante durante lo svolgimento dell'attività decisionale. La rilevazione dei segnali biometrici è necessaria al fine di trarre nuove evidenze che possano confermare o confutare le conclusioni a cui studi precedenti sono già giunti in merito ai fenomeni cerebrali e fisiologici coinvolti nelle varie attività decisionali. La raccolta dati a livello psicologico consente di tenere traccia delle percezioni del decisore e di alcune caratteristiche di questo.

2. <u>Come si svolgerà lo studio?</u>

Lo studio intende svolgere la raccolta dati nel corso di una simulazione di processo decisionale (problema del Newsvendor) e dello svolgimento di tre questionari (volti a saggiare alcune caratteristiche psicologiche del partecipante).

L'attività sperimentale prevista (problema del Newsvendor) è una serie di quattro tasks, nel corso delle quali il partecipante simula il reale contesto di un edicolante che debba decidere all'alba quante copie dell'unico giornale commercializzato in città acquistare dall'editore. Inoltre, nel corso della sessione sperimentale, il partecipante è chiamato a rispondere ai tre diversi questionari sopra menzionati.

Lo studio prevede un trattamento dei dati del tutto anonimo.

3. <u>Per quale ragione le proponiamo di partecipare?</u>

L'obiettivo principale è analizzare i segnali biometrici che caratterizzano le attività decisionali al fine di comprendere i meccanismi sottostanti tali processi.

L'ambito di ricerca risulta infatti ad oggi in fase ancora esplorativa. La proposta a partecipare è giustificata dal fatto che è necessario svolgere lo studio su volontari che siano disponibili nello svolgimento di tali attività in modo da riprodurre i diversi contesti operativi al fine di ricavare dei dati oggettivi e realistici.

4. <u>Lei è obbligato/a a partecipare allo studio?</u>

La sua partecipazione è completamente libera, il rifiuto di partecipare non comporterà alcuna conseguenza negativa. Inoltre, se dovesse cambiare idea e volesse ritirarsi dallo studio, in qualsiasi momento sarà libero/a di farlo, senza dover fornire alcuna spiegazione.

In caso di ritiro, potrà scegliere se interrompere la registrazione dei dati ma far conservare quanto raccolto fino ad allora, o se interrompere la registrazione e non far conservare quanto raccolto fino al momento del ritiro.

Nel caso i dati vengano utilizzati per lo studio, questi saranno trattati in forma del tutto anonima. Non vi sarà dunque alcun riferimento ai dati personali del partecipante, che non verranno raccolti nel corso dell'esperimento.

5. <u>Quali sono i passaggi necessari per partecipare allo studio?</u>

La partecipazione allo studio avviene previa dettagliata informazione sulle caratteristiche, sui rischi e benefici derivanti dallo stesso.

Al termine della fase informativa lei potrà acconsentire alla partecipazione allo studio firmando il modulo di consenso informato. Solo dopo che avrà espresso per iscritto il suo consenso, potrà attivamente partecipare allo studio proposto.

6. <u>Che cosa le verrà chiesto di fare?</u>

Ciascuna sessione sperimentale avrà la durata complessiva massima di un'ora e quarantacinque minuti. Nel corso della sessione sperimentale le verrà sottoposto il problema del Newsvendor (suddiviso in quattro tasks, da 10 rounds ciascuno) e le verrà chiesto di rispondere a tre diversi questionari, con l'obiettivo di raccogliere informazioni che possano migliorare la qualità dei risultati dello studio svolto.

L'intera raccolta dati sarà effettuata in forma anonima.

7. <u>Quali sono i possibili rischi ed i disagi dello studio?</u>

I disagi potrebbero essere dovuti all'applicazione di elettrodi ai polsi e agli avambracci, e di una fascia regolabile all'altezza del busto.

Gli elettrodi adesivi verrebbero utilizzati per la rilevazione dell'attività cardiaca tramite ECG e dell'attività cutanea tramite anelli regolabili per la misurazione della EDA. La fascia toracica consentirebbe invece il monitoraggio dell'attività respiratoria. Per quest'ultima non sarà richiesto al partecipante di levarsi i propri indumenti superiori.

La posizione da tenere per tutta la durata dell'esperimento è da seduto e gli saranno concessi i

movimenti minimi necessari allo svolgimento delle attività su computer, al fine di ottenere una registrazione dei segnali biometrici quanto più affidabile possibile e con il minimo numero di artefatti.Nel corso dell'intera sessione sperimentale sono previste delle brevi pause (da qualche secondo, a qualche minuto, sulla base dello specifico momento).

Gli eventi di carattere straordinario non sono correlati al compimento di questo esperimento e i ricercatori se ne deresponsabilizzano.

8. Quali sono i possibili benefici derivanti dallo studio?

Dai risultati dello studio i partecipanti potranno trarre come diretti benefici alcuni feedbacks per loro potenzialmente costruttivi. Nei giorni successivi allo svolgimento della sessione sperimentale, i partecipanti verranno infatti messi al corrente, tramite email, di alcune loro performance relative dello svolgimento del problema di gestione delle scorte (problema del Newsvendor).

Inoltre, lo studio può rappresentare un passo avanti nella ricerca nella letteratura sul Decision making in generale e può essere considerata come il punto di partenza di un processo di miglioramento delle condizioni operative in questi contesti.

9. Come viene garantita la riservatezza e sicurezza delle informazioni/dati/campioni?

Lo sperimentatore le chiederà informazioni in merito alle sue condizioni di salute (problemi cardiovascolari/respiratori). Le chiediamo questi dati perché sono strettamente necessari alla corretta esecuzione del test e alla successiva elaborazione dei dati. Non le verranno chieste altre informazioni personali.

Non verranno svolte registrazioni audio e video durante le attività sperimentali. Verranno solamente automaticamente collezionati gli orari di inserimento di tutte le risposte ai problemi, tranne quelle dell'ultimo questionario che le verrà sottoposto.

I dati biometrici rilevati e le risposte al questionario saranno associati ad un codice anonimo, perdendo la possibilità di risalire alla sua identità al termine della fase di pulizia del segnale.

Queste informazioni, così come i dati che emergeranno nel corso della ricerca, sono importanti per il corretto svolgimento dello studio. La liceità del trattamento e la riservatezza di tutte le informazioni sarà garantita secondo la normativa vigente (Regolamento europeo UE 2016/679 concernente la tutela delle persone fisiche con riguardo al trattamento dei dati personali e la libertà di circolazione di tali dati - https://www.garanteprivacy.it/regolamentoue).

1. Approvazione comitato etico

La informiamo che questo studio sarà esaminato dal Comitato Etico del *Politecnico di Torino*.

2. Altre informazioni importanti

L'originale del Consenso informato scritto da lei firmato verrà conservato dal responsabile delpresente studio, mentre lei hai diritto a riceverne una copia.

Durante lo studio, lei potrà chiedere qualsiasi informazione al Responsabile dello studio ai seguenti contatti: Professore Marco Cantamessa e Professore Giulio Zotteri.

La ringraziamo per la disponibilità.

DICHIARAZIONE DEL RESPONSABILE DELLO STUDIO

Dichiaro di aver fornito alla/al partecipante informazioni complete e spiegazioni dettagliate circa la natura, le finalità, le procedure e la durata di questo progetto di ricerca. Dichiaro, inoltre, di aver fornito alla/al partecipante il foglio informativo.

FIRMA DEL RESPONSABILE DELLO STUDIO

Data

Nome del Responsabile dello studio (*in stampatello*)

Politecnico di Torino

ESPRESSIONE DI CONSENSO INFORMATO

Io sottoscritto/a _____

DICHIARO

- di aver ricevuto spiegazioni esaurienti in merito alla richiesta di partecipazione allo studio sperimentale in oggetto e sufficienti informazioni riguardo ai rischi e ai benefici implicati nello studio, secondo quanto riportato nel foglio informativo qui allegato.
- di aver potuto discutere tali spiegazioni, di aver potuto porre tutte le domande che ho ritenutonecessarie e di aver ricevuto in merito risposte soddisfacenti;
- di essere stato, inoltre, informato del mio diritto di ritirarmi in qualsiasi momento dalla ricercastessa.

Alla luce delle informazioni che mi sono state fornite, pertanto:

☐	ACCONSENTO	☐	NON ACCONSENTO	a partecipare allo studio

LUOGO DATA

FIRMA DEL PARTECIPANTE

Bibliography

[1] P. F. Drucker, ""Management Science" and the Manager," *Management Science,* vol. 1, no. 2, p. 115, 1955.

[2] E. F. Harrison, The Managerial Decision-making Process, 5th ed., Boston, 1999.

[3] H. A. Simon, The New Science of Management Decision, 1960.

[4] D. J. McLaughlin, "Strengthening executive decision making," *Human Resource Management,* p. 443, 1995.

[5] R. D. Stacey, Strategic management and organisational dynamics: tha challenge of complexity to ways of thinking about organisations, 2007.

[6] M. Fishbein and I. Ajzen, Belief, attitude, intention and behavior: An introductionn to theory and research., 1975.

[7] I. Ajzen, "The social psychology of decision making," *Social psychology: Handbook of basic principles,* pp. 297-325, 1996.

[8] J. Orasanu and T. Connolly, "The reinvention of decision making," in *Decision making in action: Models and methods*, 1993, pp. 3-20.

[9] S. M. Narayan and S. Corcoran-Perry, "Line of reasoning as a representation of nurses' clinical," *Research in Nursing & Health, 20,* pp. 353-364, 1997.

[10] M. Porter and V. Millar, "How information gives you competitive advantage," *Harvard Business Review, 63, 4,* pp. 149-160, 1985.

[11] J. Baum and S. Wally, "STRATEGIC DECISION SPEED AND FIRM," *Strategic Management Journal, 24, 11,* pp. 1107-1129, 2003.

[12] P. A. Samuelson, "A note on the pure theory of consumer's behaviour," *Economica,* no. 5, pp. 61-71, 1938.

[13] P. A. Samuelson, Foundations of Economic Analysis, 1947.

[14] H. A. Simon, "A behavioural model of rational choice," *Quarterly Journal of Economics,* vol. 69, no. 1, p. 99, 1955.

[15] D. Kahneman, "Maps of Bounded Rationality: Psychology for Behavioral Economics," *The American Economic Review,* vol. 93, no. 5, pp. 1449-1475, 2003.

[16] E. F. Fama, "Efficient capital markets: a review of theory and empirical work," *Journal of Finance,* vol. 25, no. 2, pp. 383-417, 1970.

[17] A. Tversky and D. Kahneman, "Rational Choice and the Framing of Decisions," *The Journal of business (Chicago, Ill.),* vol. 59, no. 4, p. 252, 1986.

[18] P. Slovic, B. Fischhoff and S. Lichtenstein, Behavioral Decision Theory, 1976.

[19] M. G. Haselton, D. Nettle and P. W. Andrews, "The evolution of cognitive bias," in *The handbook of evolutionary psychology*, 2005, p. 725.

[20] M. G. Barberà-Mariné, L. Cannavacciuolo, A. Ippolito, C. Ponsiglione and G. Zollo, "The weight of organizational factors on heuristics: evidence from triage decision-making processes," *Management Decision,* vol. 57, no. 11, p. 2890, 2019.

[21] D. Arnott, "Cognitive biases and decision support systems development: a design science approach," *Information systems journal (Oxford, England), 2006, Vol.16 (1), p.55-78,* vol. 16, no. 1, p. 59, 2006.

[22] D. Arnott, "A Taxonomy of Decision Biases," 1989.

[23] G. Montibeller and D. von Winterfeldt, "Cognitive and Motivational Biases in Decision and Risk Analysis," *Risk Analysis,* vol. 35, no. 7, pp. 1230-1231, 2015.

[24] R. H. Thaler, Advances in Behavioral Finance, 1993.

[25] J. F. Yates, Judgment and Decision Making, 1990.

[26] L. R. Beach and T. Connolly, The Psychology of Decision Making: People in Organizations, 2005.

[27] J. Manoogian III, "The cognitive bias codex," Wikimedia, 6 June 2018. [Online]. Available: https://commons.wikimedia.org/wiki/File:Cognitive_bias_codex_en.svg.

[28] A. Tversky and D. Kahneman, "Judgment under uncertainty: Heuristics and biases," *Science,* vol. 185, pp. 1124-1131, 1974.

[29] H. J. Einhorn and R. M. Hogarth, "Decision Making Under Ambiguity," *The Journal of Business,* vol. 59, no. 4, pp. 225-250, 1986.

[30] C. R. Fox and J. Levav, "Familiarity bias and belief reversal in relative likelihood judgment," *Organizational Behavior and Human Decision Processes,* vol. 82, no. 2, pp. 268-292, 2000.

[31] A. Tversky and D. Kahneman, "Belief in the law of small numbers," *Psychological Bulletin,* vol. 2, pp. 105-110, 1971.

[32] D. Kahneman and A. Tversky, "Prospect Theory: An Analysis of Decisions Under Risk," *Econometrica, 47(2),* pp. 263-91, 1979.

[33] D. A. Moore and P. J. Healy, "The trouble with overconfidence," *Psychological Review,* vol. 115, no. 2, p. 502, 2008.

[34] B. E. Cushing and S. S. Ahlawat, "Mitigation of recency bias in audit judgment: the effect of documentation," *Auditing,* vol. 15, no. 2, pp. 110-122, 1996.

[35] D. Kahneman and A. Tversky, "On the psychology of prediction," *Psychological Review,* vol. 80, no. 4, pp. 237-251, 1973.

[36] J. Shaanan, "Sunk costs and resource mobility: an empirical study," *Review of Industrial Organization,* vol. 9, no. 6, pp. 717-730, 1994.

[37] C. R. Carter, L. Kaufmann and A. Michel, "Behavioral supply management: a taxonomy of judgment and decision-making biases," *International Journal of Physical Distribution & Logistics Management,* vol. 37, no. 8, pp. 631-669, 2007.

[38] W. E. Remus and J. E. Kottemann, "Toward Intelligent Decision Support Systems: An Artificially Intelligent Statistician," *MIS Quarterly,* vol. 10, no. 4, pp. 403-418, 1986.

[39] S. Gill, J. Stockard, M. Johnson and S. Williams, "Measuring gender differences: The expressive dimension and critique of androgyny scales," *Sex Roles, 17,* pp. 375-400, 1987.

[40] J. T. Wood, Gendered lives: Communication, gender, and culture, 1990.

[41] W. Gardner, D. Scherer and M. Tester, "Asserting scientific authority: Cognitive development and adolescent legal rights," *American Psychologist, 44,* pp. 895-902, 1989.

[42] I. E. Dror, M. Katona and K. Mungur, "Age difference in decision making: To take a risk or not?," *Gerontology, 44,* pp. 67-71, 1998.

[43] Y. Chen and Y. Sun, "Age differences in financial decision-making: Using simple heuristics," *Educational Gerontology, 29,* pp. 627-635, 2003.

[44] D. Moshman, "Adolescent reasoning and adolescent right," *Human Development, 36,* pp. 27-40, 1993.

[45] M. L. Sanz de Acedo Lizárraga, M. T. Sanz de Acedo Baquedano and M. Cardelle-Elawar, "Factors that affect decision making: gender and age differences," *International Journal of Psychology and Psychological Therapy, vol. 7, num. 3,* pp. 381-391, 2007.

[46] P. T. Costa and R. R. McCrae, "Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI)," *Psychological Assessment Resources,* 1992.

[47] J. Strohhecker and A. Größler, "Do personal traits influence inventory management performance?—The case of intelligence, personality, interest and knowledge," *International journal of production economics, Vol.142 (1),* pp. 37-50, 2013.

[48] J. L. Holland, "A personality inventory employing occupational titles," *Journal of Applied Psychology, 42,* p. 336–342, 1958.

[49] J. L. Holland, "Making Vocational Choices : A Theory of Vocational Personalities and Work Environments, 3 ed.," *Psychological Assessment Resources,* 1997.

[50] J. F. Yates and G. A. Potorowski, "Evidence-based decision management," in *The Oxford handbook of evidence-based management*, vol. 136, 2012, pp. 198-222.

[51] Y. E. Chentsova-Dutton and A. Vaughn, "Let me tell you what to do: Cultural differences in advice-giving," *Journal of Cross-Cultural Psychology, 43 (5),* p. 687–703, 2012.

[52] J. F. Yates and S. de Oliveira, "Culture and decision making," *Organizational behavior and human decision processes,* vol. 136, pp. 106-118, 2016.

[53] D. Kahneman, Thinking, Fast and Slow, 2011.

[54] K. E. Stanovich and R. F. West, "Individual Differences in Reasoning: Implications for the Rationality Debate," *Behavioral and Brain Sciences, 23(5),* pp. 645-65, 2000.

[55] A. Shleifer, "Psychologists at the Gate: A Review of Daniel Kahneman's "Thinking, Fast and Slow"," *Journal of Economic Literature,* vol. 50, no. 4, pp. 1080-1091, 2012.

[56] P. L. Ackerman, "A theory of adult intellectual development: process, personality, interests, and knowledge," *Intelligence 22,* p. 227–257, 1996.

[57] C. Achar, J. So, N. Agrawal and A. Duhachek, "What we feel and why we buy: the influence of emotions on consumer decision-making," *Current opinion in psychology,* vol. 10, pp. 166-170, 2016.

[58] H. Selye, "A syndrome produced by diverse nocuous agents," *Nature, 138:32,* 1936.

[59] L. H. Berghorst, R. Bogdan, M. J. Frank and D. A. Pizzagalli, "Acute stress selectively reduces reward sensitivity," *Frontiers in human neuroscience,* vol. 7, pp. 133-133, 2013.

[60] R. Bogdan and D. A. Pizzagalli, "Acute Stress Reduces Reward Responsiveness: Implications for Depression," *Biological Psychiatry,* vol. 60, pp. 1147-1154, 2006.

[61] L. Schwabe and O. T. Wolf, "Stress prompts habit behavior in humans," *The Journal of neuroscience, 2009, Vol.29 (22),* pp. 7191-7198, 2009.

[62] A. L. Krain, A. M. Wilson, R. Arbuckle, F. X. Castellanos and M. P. Milham, "Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making," *NeuroImage, Vol.32 (1),* pp. 477-484, 2006.

[63] M. Hsu, M. Bhatt, R. Adolphs, D. Tranel and C. Camerer, "Neural Systems Responding to Degrees of Uncertainty in Human Decision-Making," *Science (American Association for the Advancement of Science), Vol.310 (5754),* pp. 1680-1683, 2005.

[64] M. L. Platt and S. A. Huettel, "Risky business: the neuroeconomics of decision making under uncertainty," *Nature Neuroscience, vol. 11,* p. 398–403, 2008.

[65] A. J. Porcelli and M. R. Delgado, "Stress and decision making: effects on valuation,," *Current opinion in behavioral sciences,* vol. 14, pp. 33-39, 2017.

[66] R. M. Hogarth, C. R. McKenzie, B. J. Gibbs and M. A. Marquis, "Learning to make decisions: When incentives help and hinder," 1989.

[67] J. A. Easterbrook, "The effect of emotion on cue utilization and the organization of behavior," *Psychological Review, 66 (3),* pp. 183-201, 1959.

[68] D. F. Dinges, F. Pack, K. Williams, K. A. Gillen, J. W. Powell, G. E. Ott, A. I. Pack and C. Aptowicz, "Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night," *Sleep, 20,* pp. 267-277, 1997.

[69] J. J. Pilcher and A. I. Huffcutt, "Effects of sleep deprivation on performance: A meta-analysis," *Sleep, 19(4),* pp. 318-326, 1996.

[70] W. D. S. Killgore, T. J. Balkin and N. J. Wesensten, "Impaired decision making following 49 h of sleep deprivation," *Journal of Sleep Research, 15,* pp. 7-13, 2006.

[71] B. S. Mckenna, D. L. Dickinson, H. J. Orff and S. P. A. Drummond, "The effects of one night of sleep deprivation on known-risk and ambiguous-risk decisions," *Journal of Sleep Research, Vol. 16, Issue 3,* vol. 16 , pp. 245-252, 2007.

[72] D. L. Dickinson and T. McElroy, "Bayesian versus heuristic-based choice under sleep restriction and suboptimal times of day," *Games and Economic Behavior,* vol. 115, pp. 48-59, 2019.

[73] D. L. Dickinson and T. McElroy, "Sleep restriction and circadian effects on social decisions," *European Economic Review,* vol. 97, pp. 57-71, 2017.

[74] J. B. Holbein, J. P. Schafer and D. L. Dickinson, "Insufficient sleep reduces voting and other prosocial behaviors," *Nature Human Behaviour,* vol. 3, pp. 492-500, 2019.

[75] B. M. Hood and D. Bruck, "Information processing in narcolepsy," *Journal od Sleep Research,* vol. 5, pp. 128-134, 1997.

[76] R. Wimmer, R. P. Hoffman, R. A. Bonato and A. R. Moffitt, "The effects of sleep deprivation on divergent thinking and attention processes," *Journal of Sleep Research, Vol. 1, Issue 4,* pp. 223-230, 1992.

[77] Y. Harrison and J. A. Home, "One Night of Sleep Loss Impairs Innovative Thinking and Flexible Decision Making," *Organizational Behavior and Human Decision Processes, Vol. 78, Issue 2,* pp. 128-145, 1999.

[78] P. Y. Martin and S. Marrington, "Morningness–eveningness orientation, optimal time-of-day and attitude change: Evidence for the systematic processing of a persuasive communication," *Personality and Individual Differences, Vol. 39, Issue 2,* pp. 367-377, 2005.

[79] T. H. Monk and V. C. Leng, "Interactions Between Inter-Individual and Inter-Task Differences in the Diurnal Variation of Human Performance," *The Journal of Biological and Medical Rhythm Research, Vol. 3, Issue 3,* pp. 171-177, 1986.

[80] E. Howarth and M. S. Hoffman, "A multidimensional approach to the relationship between mood and weather," *British Journal of Psychology,* vol. 75, no. 1, pp. 15-23, 1984.

[81] H. A. Simon, "Motivational and emotional controls of cognition," *Psychological Review,* vol. 74, pp. 23-39, 1967.

[82] G. R. J. Hockey, A. John Maule, P. J. Clough and L. Bdzola, "Effects of negative mood states on risk in everyday decision making," *Cognition and Emotion,* vol. 14, no. 6, pp. 823-855, 2000.

[83] P. M. Parker and N. T. Tavassoli, "Homeostasis and Consumer Behaviour across Cultures," *International Journal of Research in Marketing,* vol. 17, no. 1, pp. 33-53, 2000.

[84] F. W. Schneider, W. A. Lesko and W. A. Garrett, "Helping behaviour in hot, comfortable and cold temperature: A field study," *Environment and Behaviour,* vol. 12, no. 2, pp. 231-240, 1980.

[85] M. C. Keller, B. L. Fredrickson, O. Ybarra, S. Côté, K. Johnson, J. Mikels, A. Conway and T. Wager, "A warm heart and a clear head: The contingent effects of weather on mood and cognition," *Psychological Science,* vol. 16, no. 9, pp. 724-731, 2005.

[86] C. P. May and L. Hasher, "L. Synchrony effects in inhibitory control over thought and action," *J. Exp. Psychol. Hum. Percept. Perform.,* vol. 24, p. 363–379, 1998.

[87] K. K. Ingram, A. Ay, S. B. Kwon, K. Woods, S. Escobar, M. Gordon, I. H. Smith, N. Bearden, A. Filipowicz and K. Jain, "Molecular insights into chronotype and time-of-day effects on decision-making," *Scientific reports, 2016,* vol. 6, no. 1, pp. 29392-29392, 2016.

[88] J. Li, P. L. Pearce, A. M. Morrison and B. Wu, "Up in Smoke? The Impact of Smog on Risk Perception and Satisfaction of International Tourists in Beijing," *The international journal of tourism research,* vol. 18, no. 4, pp. 373-386, 2016.

[89] B. Zhang, B. Wu and J. Liu, "PM2.5 pollution-related health effects and willingness to pay for improved air quality: Evidence from China's prefecture-level cities," *Journal of cleaner production,* vol. 273, p. 122876, 2020.

[90] K. Zhang, Y. Hou, G. Li and Y. Huang, "Tourists and Air Pollution: How and Why Air Pollution Magnifies Tourists' Suspicion of Service Providers," *Journal of Travel Research,* vol. 59, no. 4, pp. 661-673, 2020.

[91] X. Zhang, Z. Ni, Y. Wang, S. Chen and B. Xia, "Public perception and preferences of small urban green infrastructures: A case study in Guangzhou, China," *Urban Forestry & Urban Greening,* vol. 53, p. 126700, 2020.

[92] J. H. Wang, J. H. Kim, J. H. Moon and H. J. Song, "The Effect of Smog-Related Factors on Korean Domestic Tourists' Decision-Making Process," *International journal of environmental research and public health,* vol. 17, no. 10, p. 3706, 2020.

[93] M. E. Schweitzer and G. P. Cachon, "Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence," *Management Science,* vol. 46, no. 3, pp. 404-420, 2000.

[94] F. Knight, Risk, uncertainty and profit., Boston, NY, 1921.

[95] S. Sniazhko, "Uncertainty in decision-making: A review of the international business literature," *Cogent Business & Management, ,* vol. 6, no. 1, 2019.

[96] A. Tversky and W. Edwards, "Information versus reward in binary choices," *Journal of Experimental Psychology,* vol. 71, no. 5, pp. 680-683, 1966.

[97] J. C. McCullers, "Issues in learning and motivation," in *The hidden costs of reward*, 1978, pp. 5-18.

[98] J. W. Payne, J. R. Bettman and E. J. Johnson, "Adaptive strategy selection in decision making: Learning, Memory & Cognition," *Journal of Experimental Psychology,* vol. 14, p. 534–552, 1988.

[99] N. H. Lurie and J. M. Swaminathan, "Is timely information always better? The effect of feedback frequency on decision making," *Organizational behavior and human decision processes,* vol. 108, no. 2, pp. 315-329, 2009.

[100] R. M. Hogarth, B. J. Gibbs, C. R. McKenzie and M. A. Marquis, "Learning from feedback: Exactingness and incentives," *Journal of Experimental Psychology: Learning, Memory & Cognition,* vol. 17, p. 734–752., 1991.

[101] J. G. March, "Exploration and exploitation in organizational learning," *Organization Science,* vol. 2, pp. 71-87, 1999.

[102] C. Camerer, "Progress in behavioral game theory," *Journal of Economic Perspectives,* vol. 11, p. 167–188, 1997.

[103] A. N. Kluger and A. DeNisi, "The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory," *Psychological Bulletin,* vol. 119, pp. 254-284, 1996.

[104] J. A. F. Stoner, *A comparison of individual and group decisions involving risk (Doctoral dissertation, Massachusetts Institute of Technology),* 1961.

[105] J. P. Guilford, Fundamental statistics in psychology and education, 1950.

[106] J. P. Guilford, The nature of human intelligence., 1967.

[107] G. Gigerenzer and W. Gaissmaier, "Decision Making: Nonrational Theories," *International Encyclopedia of the Social & Behavioral Sciences, 2nd edition,* vol. 5, 2015.

[108] P. Guo and Y. Li, "Approaches to multistage one-shot decision making," *European Journal of Operational Research,* vol. 236, no. 2, pp. 612-623, 2014.

[109] W. Edwards, "Dynamic decision theory and probabilistic information processing," *The Journal of Human Factors and Ergonomics Society,* vol. 4, pp. 59-73, 1962.

[110] A. A. Bostian, C. A. Holth and A. M. Smith, "Newsvendor "Pull-to-Center" Effect: Adaptive Learning in a Laboratory Experiment," *Manufacturing & Service Operations Management,* vol. 10, no. 4, pp. 590-608, 2008.

[111] S. Yamini, "Behavioral perspective of newsvendor ordering decisions: review, analysis and insights," *Management Decision,* vol. 59, no. 2, pp. 240-257, 2021.

[112] G. E. Bolton and E. Katok, "Learning by Doing in the Newsvendor Problem: A Laboratory Investigation of the Role of Experience and Feedback," *Manufacturing & Service Operations Management,* vol. 10, no. 3, pp. 519-538, 2008.

[113] N. C. D. Truong, X. Wang, H. Wanniarachchi, Y. Lang, S. Nerur, K. Chen and H. Liu, "Mapping and understanding of correlated electroencephalogram (EEG) responses to the newsvendor problem," *Scientific reports,* vol. 12, no. 1, pp. 13800-13800, 2022.

[114] K. J. Arrow, T. Harris and J. Marschak, "Optimal Inventory Policy," *Econometrica,* vol. 19, no. 3, pp. 250-272, 1951.

[115] T. M. Whitin, "Inventory control and price theory," *Management Science,* vol. 2, pp. 61-80, 1955.

[116] R. Croson and K. Donohue, "Behavioral Causes of the Bullwhip Effect and the Observed Value of Inventory Information," *Management Science,* vol. 52, no. 3, pp. 323-336, 2006.

[117] J. Forrester, "Industrial dynamics: A major breakthrough for decision makers," *Harvard Business,* vol. 36, pp. 37-66, 1958.

[118] U. Benzion, Y. Cohen, R. Peled and T. Shavit, "Decision-Making and the Newsvendor Problem: An Experimental Study," *The Journal of the Operational Research Society,* vol. 59, no. 9, pp. 1281-1287, 2008.

[119] Y. Ren and R. Croson, "Overconfidence in Newsvendor Orders: An Experimental Study," *Management science,* vol. 59, no. 11, pp. 2502-2517, 2013.

[120] K. L. Schultz, L. W. Robinson, L. J. Thomas, J. Schultz and J. O. McClain, "The use of framing in inventory deicisions," *Production and operations management,* vol. 27, no. 1, pp. 49-57, 2018.

[121] G. Zotteri, M. Del Nibletto and S. Colombo, "Experimental Economics and the Newsvendor Problem," Under review.

[122] B. B. Moritz, A. V. Hill and K. L. Donohue, "Individual differences in the newsvendor problem: Behavior and cognitive reflection," *Journal of operations management,* vol. 31, no. 1-2, pp. 72-85, 2013.

[123] S. Frederick, "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives,* vol. 19, no. 4, pp. 25-42, 2005.

[124] P. Brandimarte and G. Zotteri, Introduction to distribution logistics, John Wiley & Sons, 2007.

[125] Y. Ren, D. C. Croson and R. T. A. Croson, "The overconfident newsvendor," *Journal of the Operational Research Society,* vol. 68, pp. 496-506, 2017.

[126] E. A. Crone, R. J. M. Somsen, B. V. Beek and M. W. Van Der Molen, "Heart rate and skin conductance analysis of antecendents and consequences of decision making," *Psychophysiology,* vol. 41, no. 4, pp. 531-540, 2004.

[127] M. De Couck, R. Caers, L. Musch, J. Fliegauf, A. Giangreco and Y. Gidron, "How breathing can help you make better decisions: Two studies on the effects of breathing patterns on heart rate variability and decision-making in business cases," *International Journal of Psychophysiology,* vol. 139, pp. 1-9, 2019.

[128] C. Primi, K. Morsanyi, F. Chiesi, M. A. Donati and J. Hamilton, "The Development and Testing of a New Version of the Cognitive Reflection Test Applying Item Response Theory (IRT)," *Journal of behavioral decision making,* vol. 29, no. 5, pp. 453-469, 2016.

[129] A. L. Sachs and S. Minner, "The data-driven newsvendor with censored demand observations," *International journal of production economics,* vol. 149, pp. 28-36, 2014.

[130] W. K. Balzer, M. E. Doherty and R. J. O'Connor, "Effects of cognitive feedback on performance," *Psychological Bulletin,* vol. 106, p. 410–433, 1989.