
UNIVERSITY OF TRENTO

DEPARTMENT OF PSYCHOLOGY AND COGNITIVE SCIENCES

DOCTORAL SCHOOL IN

PSYCHOLOGICAL SCIENCES AND EDUCATION

XXVIII CYCLE


**Doctoral Dissertation**

---

# Measuring real-time dynamics underlying online rating processes:

# Models, methods, applications

---


*Phd Scholar*
Antonio Calcagnì

*Advisor*
prof. Luigi Lombardi

*Head of the doctoral school*
prof.ssa Paola Venuti



December 2015

*"A stento immaginiamo le cose della terra, scopriamo con fatica quelle a portata di mano; ma chi ha investigato le cose del cielo? Chi avrebbe conosciuto il tuo volere, se tu non gli avessi dato la sapienza e dall'alto non gli avessi inviato il tuo santo spirito? Così vennero raddrizzati i sentieri di chi è sulla terra; gli uomini furono istruiti in ciò che ti è gradito [...]."*

Sapienza 9, 16-18

Come ogni frutto della creatività umana non può originarsi da un lavoro solingo ma si tesse e ritesse all'interno di una tela articolata dalle relazioni umane, anche questo elaborato emerge dall'incontro fruttuoso con altre persone che coi loro saperi hanno collaborato a rendere il tutto un composito omogeneo e strutturato al suo interno. È perciò atto piacevolmente dovuto quello di esprimere la mia gratitudine verso coloro che mi hanno guidato ed accompagnato in questo viaggio volto ad approfondire una piccola parte del vasto sapere scientifico, all'interno di un progressivo percorso di studio e conoscenza delle pratiche che la concorrono a formare.

Nella consapevolezza che le parole sono insufficienti ad esprimere i più profondi vissuti, un primo sincero ringraziamento è rivolto al prof. Luigi Lombardi che con paziente e costante presenza mi ha traghettato verso l'apprendimento di quegli utensili necessari alla buona pratica della ricerca scientifica. Come un vasaio fa con l'argilla, ha progressivamente favorito la mia crescita scientifica e personale sino alla maturazione oggi raggiunta, entro un processo che ancora prosegue e proseguirà. Come un "operaio" ho appreso quest'umile mestiere presso la sua "bottega" di esperienza e saggezza. La mia attuale maturazione è primariamente frutto dell'intenso lavoro fatto con lui in questi anni.

Un secondo ringraziamento è per coloro che più avanti in questo percorso hanno saputo donarmi il loro tempo ma anche la loro amicizia. In particolare, Lorenzo, Simone e Mara che nelle diversità dei loro carismi hanno permesso la realizzazione di alcuni progetti insieme. Sentimento di gratitudine anche per i miei amici e colleghi Francesco, Janet, Andrea, Yagmur e Mauro insieme ai quali ho condiviso questo percorso.

Con viva gratitudine non posso non ricordare tutti i docenti di questa Scuola che hanno donato il loro tempo nell'organizzare ed incoraggiare le attività di formazione. In particolare, la prof.ssa Paola Venuti che in qualità di coordinatrice ci ha guidato in questo percorso di crescita personale e professionale.

Un ultimo ringraziamento è per Martina che con la sua saggezza di donna mi ha accompagnato nelle fasi finali di questo percorso. Coniugando cuore e ragione con la dovuta semplicità dell'abbandono provvidenziale, mi ha supportato ed incoraggiato amorevolmente nell'affrontare le fatiche del crescere e del maturare.

# Contents

# Introduction

This thesis collects a set of research articles and working papers produced during my doctoral program. Overall, they refer to:

(i) the general problem of collecting measures for the dynamic processes underlying human rating tasks (section *Models, methods, instruments*)

(ii) the proposal of new statistical methods to analyse data structures arising from the dynamic rating measurements (section *Statistical techniques*).

The section *Models, methods, instruments* contains three chapters on models and methods designed to represent the dynamics of rating evaluations together with some relevant applications. In particular, the latter refer to two different contexts, namely Likert-type questionnaires (Chapters 1 and 2) and two-choice decision making (Chapter 3). Finally, the section also describes two new measurement instruments for dynamic ratings, DYFRAT (Chapter 1) and DYFRAS (Chapter 2). In a similar way, the section *Statistical techniques* contains three works on novel statistical methodologies to run data analysis in the context of "structured data", namely fuzzy numbers (Chapters 4 and 5) and interval-valued data (Chapter 6). The justification for structured-data comes directly from the treatment of the dynamic rating components in the sense that such a kind of data can offer a simple and flexible formal representation to model the information related to the rating properties.

Traditionally, rating objects like attitudes, opinions, and subjective evaluations have been investigated by means of Likert-type questionnaires

and/or similar methodologies (e.g., see: Aiken, 1996). In this context, the extracted information have been usually codified in terms of simple and discrete numerical representations (i.e., final discrete responses and eventually their associated response times). The standard assumption here is that these standard observable measures are "end products" of the entire process of rating and, as such, can be used as indicators of the raters' overall performance. In this way, the dynamics underlying final rating measures are clearly not detected with the consequence of loss of information. Although this issue is not completely new in the psychological tradition (e.g., see: Anderson, 1981; Tourangeau and Rasinski, 1988; Smith and DeCoster, 2000; Link, Townsend, and Ashby, 1985), over the recent years it is becoming popular again (e.g., see: Rosenbaum, 2005; Dale, Hindy, and Spivey, 2006). This can be also noticed in other scientific traditions (e.g., see the recent methods employed in neuroscientific investigations such as the GPFA method or the so-called trial-by-trial analysis - Byron et al., 2009; Delorme and Makeig, 2004). In this respect, studying the dynamics at the base of the rating mechanisms could give us a pragmatical but still useful window on how human evaluations, opinion, and attitude may arise.

It should be noted that several models and methods might be used in studying underlying rating processes. For instance, one could use probabilistic modelisations such as dynamic factor analysis, hidden Markov models, multinomial trees modeling, and IRT based models. However, for the sake of simplicity and convenience, we preferred to approach this relevant topic from a descriptive "metrological perspective". More precisely, here we refer to the study of empirical data as they appear in their low level representation, to the choice of the basic model to represent them and the related measurement systems. Put in other terms, we would prefer to address the main question - how empirical rating data can be modeled in order to extract from them as much information as possible - before any further statistical manipulation. This would focus on the basic and important problem of the "first data representation". However, in doing so, we

do not want to embed our questions within the historical debate between psychometrics and mathematical psychology about the nature of the psychological measurement (e.g., see: Luce, 1996; Sijtsma, 2012; Ratcliff, 1998; Michell, 1999). Although we know that theoretical and methodological problems could arise in this context, we discuss our findings by adopting an "exploratory attitude", considering our proposal as first attempts to light up some issues that are related with the modelisation of real-time rating processes.

# Models, methods, instruments

# Chapter 1

# Dynamic Fuzzy Rating Tracker (DYFRAT): a novel methodology for modeling real time dynamic cognitive processes in rating scales

The content of the chapter has been previously published as: Calcagnì, A., & Lombardi, L. (2014). Dynamic Fuzzy Rating Tracker (DYFRAT): a novel methodology for modeling real-time dynamic cognitive processes in rating scales. *Applied soft computing*, 24, 948-961.

## 1.1 Introduction

In many empirical research areas such as, for instance, psychology, sociology, organizational and management sciences, marketing, and epidemiology, *rating scales* represent a widely used, simple and flexible tool for measuring attitudes, opinions, and subjective preferences (Göb, McCollin, and Ramalhoto, 2007; Miller and Salkind, 2002; Aiken, 1996; Pettit, 2002). Let us assume that we are interested in measuring a person's degree of happiness. We could do this in a number of different ways, but one direct and efficient approach would be simply to ask the person, 'How happy are you?' and require them evaluate themselves on a Likert-type rating scale, ranging from 'very unhappy' to 'very happy'. Rating scales typically consists of a variable to be measured (e.g., 'happiness') and a set of *anchor points* from which the rater selects the most appropriate description (e.g., very unhappy, moderately unhappy, neither, moderately happy,

very happy). One widely used type of rating scale is the so-called *numerical* scale, where the anchor points either explicitly or implicitly are defined numerically (e.g., 1: low, 2: average, 3: high). Like checklists, rating scales are used for a wide variety of assessment purposes. For example, rating scales can be used to have one individual evaluate another, for example, when a physician might asses a patient as to degree of obesity, but the rating scales can also be applied as self-report measures. Unlike other types of ratings, self-report scales require the person provide a direct and explicit rating of their own behavior/opinion/preference, etc. Of course, the main assumption behind self-report measures is that individuals are in the best position to report their own opinion in a direct and transparent way. The great diffusion and success of rating scales are mainly due to the following major reasons that are all well documented in the literature: (1) rating scales can be administered to large groups of respondents at one single setting; (2) they can be administered under conditions that guarantee anonymity; (3) they allow the rater to proceed at their own pace; (4) they present uniformity of procedure; (5) they allow for great flexibility - for example, take-home questionnaires; and (6) the results are more amenable to statistical analyses (in particular for numerical scales) (Domino and Domino, 2006).

However, over the years several criticisms have been arisen against the empirical validity of rating scales. For example, because of the discrete and crisp nature of rating scales, some raters tend to avoid extreme categories in the anchor points (central tendency or restriction of range problem) (Saal, Downey, and Lahey, 1980). In general, we may think of human opinions and subjective preferences as being characterized by more vague and imprecise information than the ones actually described by standard rating scales. Moreover, in some circumstances, the honesty assumption implicitly assumed in self-report rating administrations appears to be simply unrealistic. So, for example, in personnel selection surveys some candidates might not be so candid in admitting their imperfections and,

therefore, the associated self-report evaluations could be definitively biased (e.g., Furnham, 1986). Finally, another important limitation pertains the idea that what is being reported using rating evaluations is something that allows us to objectively explain and describe a person's behavior. However, several psychologists argue that such assumption is, in principle, ill-posed and that what we really need, instead, is to directly focus on the observable behavior involved in the process of rating (e.g., Golfried and Kent, 1972).

In order to overcome the limitations of standard rating scales, some researchers have applied fuzzy set theory (FST) to directly modeling imprecise features of human rating evaluations (Sáa et al., 2013; Li, 2012). In particular, in the rating scale problem, FST has been mainly used in two different contexts: i) for modeling data obtained by means of standard rating scales (*fuzzy conversion scales*) or ii) for directly quantifying empirical evaluations using *fuzzy rating scales*. In the first case, FST is applied a-posteriori as a procedure for converting standard rating data into fuzzy data (i.e., raters express their judgements using a traditional rating scale which is subsequently converted into a fuzzy structure). By contrast, in the second case, FST is applied a-priori as a general interface for directly capturing fuzzy rating data (i.e. raters give their evaluations by means of computerized tools that allow to directly use fuzzy sets in place of crisp numbers).

In line with these approaches, in the present contribution we propose a novel methodology, called DYFRAT (Dynamic Fuzzy Rating Tracker), to measure some relevant behavioral dynamics involved in the rating process. Likewise fuzzy rating scales and fuzzy conversion scales, also our tool represents human rating evaluations in terms of fuzzy sets. However, unlike fuzzy conversion and fuzzy rating scales, DYFRAT captures (in addition) some physical and biometric characteristics underlying observable behaviors involved in the process of rating and considers fuzziness as a natural property that spontaneously arises from the observed data. In this

respect, DYFRAT is a formal procedure that explicitly focuses on the behavioral dynamics of rating and provides a continuous on-line measure of the cognitive aspects involved in this process.
The remainder of this chapter is organized as follows. In the second section we describe a comprehensive survey of the state-of-the-art of FST applications in human rating problems. In the third and fourth sections we present our new methodology. In the fifth section we describe a first computerized implementation of DYFRAT system. In the sixth section we show some empirical applications of DYFRAT to real data, whereas in the seventh section we conclude this chapter by providing some final comments.

## 1.2 Currently used methods for fuzzy ratings

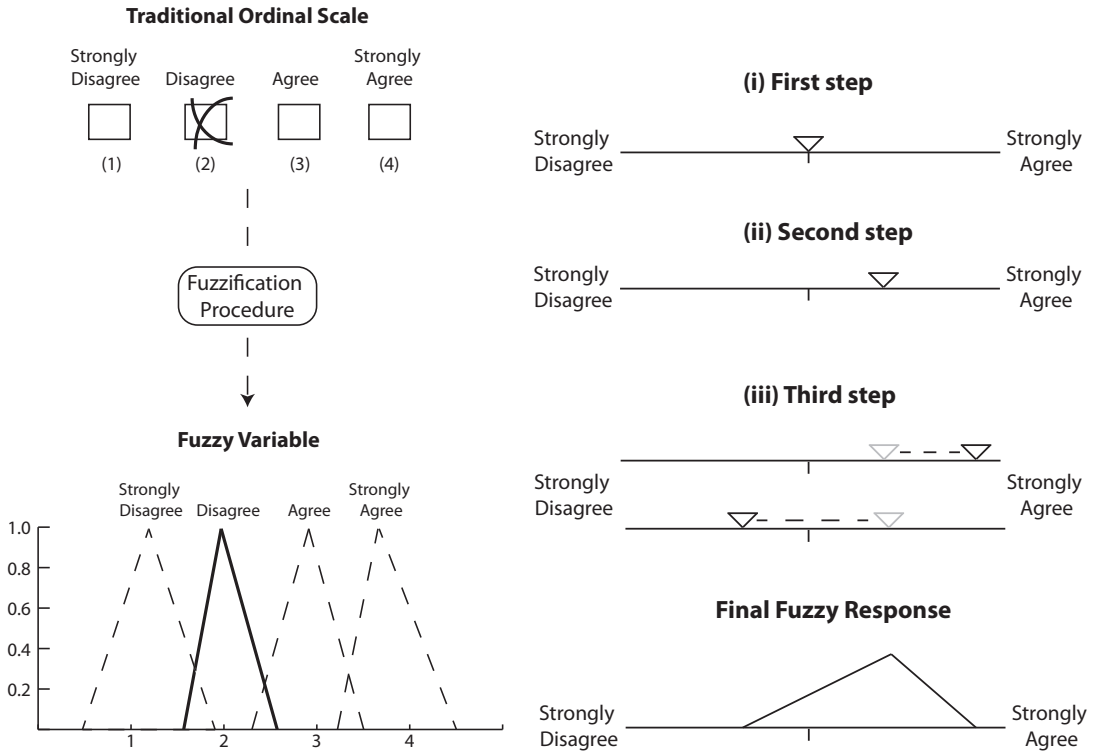
### 1.2.1 Fuzzy conversion scales

Fuzzy conversion scales (FCS) are computational procedures based on a fuzzy system which convert standard crisp rating data into a set of fuzzy data. Figure 1.1a shows a graphical representation of the rationale underlying the FCS approach. In general, two perspectives can be adopted to derive the conversion scale procedure from crisp rating data: i) an expert-based approach and ii) an empirically-based approach. In the first perspective, a researcher (a-priori) sets the main features of the fuzzy variables (the number and shape of fuzzy sets, their overlapping degrees, etc.) as well as some properties of the fuzzy system used in the conversion process (input and output space definition, implication rules, implication surface, etc). In the second approach, the fuzzy variables, are obtained using suitable statistical or data-mining procedures applied on the original crisp data (e.g., fuzzy clustering, histogram based methods, probability-possibility transformations).
FCS has found wide application in different empirical contexts. For instance, Chang and Chung used FCS to study service quality of domestic

airline companies (Chang and Yeh, 2002), Chan and Kao analysed customers'satisfaction (Chan, Kao, and Wu, 1999), Benítez et al. proposed a FST based method to measure hotels quality (Benítez, Martín, and Román, 2007), Lalla et al. developed a system for teaching activity evaluations (Lalla, Facchinetti, and Mastroleo, 2005), Costa et al. proposed an application for perception of disability evaluation (Costa et al., 2004), Li developed a novel type of fuzzy Likert scale (Li, 2012), Moon et al. described a system for ranking candidates in military organizations (Moon, Lee, and Lim, 2010), and Chen et al. proposed a fuzzy approach for selecting environment-watershed plans (Chen et al., 2011).
However, fuzzy conversion scales are also characterized by some limitations which are related to their general definition of fuzzy variables. For example, aspects such as the number or type of fuzzy sets and the amount of overlap between consecutive fuzzy sets in the scale, represent all sensitive steps in the scale construction process. Moreover, the realization of the conversion procedure, usually implemented by means of specific fuzzy systems, is generally not a trivial task as it requires several delicate phases (e.g., evaluation of the uniqueness of the final fuzzy representation, measurement of the overall goodness-of-fit). In general, although the adoption of the empirically-based methodology may overcome some of the drawbacks of the expert-based approach (Ciavolino, Salvatore, and Calcagnì, 2013; Medasani, Kim, and Krishnapuram, 1998), FSC still inherits some of the limitations of standard rating scales, such as for example, the problem of selecting extreme categories in the scale and, more importantly, the fact that this methodology cannot provide any information about the underlying dynamics of rating.

### 1.2.2 Fuzzy rating scales

A Fuzzy rating scale (FRS) is a graphical tool which allows to collect human rating data by means of a computerized procedure entirely based on the FST rationale. Generally, the FRS procedure permits to directly draw, on a continuous scale with a movable cursor (or mouse pointer), a fuzzy




(A) *General diagram for FCS* (B) *General diagram for FRS*

FIGURE 1.1: Graphical comparison between FCS and FRS

set which would then represent the subjective judgement or evaluation of the rater. A two-step paradigm is usually assumed in FRS applications. First, the respondent chooses a point on a continuous scale as h(er/is) preferred position (Fig. 1.1b-i). Next, to express how fuzzy the rater feels about the original preference, s(he) draws a line to the right (Fig. 1.1b-ii) and another one to the left (Fig. 1.1b-iii). In an alternative version of the FRS procedure, a triangular or trapezoidal fuzzy set lying on a continuous line can be directly used, as a movable cursor, by the rater who subsequently can modify it by expanding or contracting its sides in order to express the response uncertainty. Also FRS has been successfully applied in many real contexts. For instance, Hesketh and Hesketh proposed, for the first time, a measurement tool for subjective attitudes called computerized fuzzy rating scale (Hesketh et al., 1988; Hesketh, Pryor, and Hesketh, 1988; Hesketh and Hesketh, 1994). Costas et al. applied fuzzy rating scales for

the evaluation of psychological constructs (Costas, Maranon, and Cabrera, 1994), finally, Gil et al. described a free fuzzy-numbered response format questionnaire and compared it with standard rating scales (Gil and González-Rodríguez, 2012).
However, the application of fuzzy scaling methods can also be problematic. In particular, FRS can lack some ecological validity as subjects must follow specific training periods for learning how the instrument exactly works, what the concept of fuzziness is and, above all, how it can be used to directly express their judgements or evaluations. Moreover, FRS can also be criticized because this technique asks respondents to provide information that may seem difficult or even tricky to them. For this reason, the basic assumption underpinning FRS, namely that human raters are able to explicitly express the fuzziness of their judgements or evaluations by means of fuzzy sets, could be questionable. Therefore, due to the high complexity of the rating procedure, we cannot exclude that in some situations data collected using the FRS framework could be distorted or even meaningless (Howard and Dailey, 1979; Furnham and Henderson, 1982; Windschitl and Wells, 1996).

## 1.3 Dynamic Fuzzy Rating Tracker: theory

The traditional rating scale paradigm often regards human rating as a discrete-stage process in which the final response represents its final stage. Unfortunately, the observed final response captures only the outcome of the rating process while the real-time cognitive dynamics that occur during this process are lost. To overcome this relevant limitation, we propose a new methodology (DYFRAT) which is designed to track real-time mental processes by using the so-called Mouse Tracking Methodology (MTM). This methodology allows to record peculiar cognitive information derived from the motor control of the computer-mouse (or other pointer devices such as, for example, trackball, joystick, light pen, wii system,

laser pen, etc.). Just as saccadic eye movements, brain waves, pupil diameters, reaction-times, also computer-mouse movements can provide a continuous on-line measure of the cognitive processes involved in human ratings (Cacioppo and Tassinary, 1990; Franco-Watkins and Johnson, 2011). Moreover, several studies have shown as the computer-mouse can be considered as a reliable, cheap, non intrusive and accurate data acquisition device that can successfully be used in cognitive as well as behavioral research (Jansen, Blackwell, and Marriott, 2003; Hwang et al., 2005; Chen, Anderson, and Sohn, 2001; Mueller and Lockerd, 2001; Freeman and Ambady, 2010; O'Reilly and Plamondon, 2011). In the DYFRAT approach, the process of human rating is described as a temporal and dynamic changing course of data information (Morein-Zamir et al., 2006; Johnson et al., 2012). More precisely, because subjective judgements or evaluations can be conceived as the output of peculiar *underlying cognitive processes* (Morein-Zamir et al., 2006; Magnuson, 2005; Freeman, Dale, and Farmer, 2011), DYFRAT uses some biometric physical measures (which are associated with the rating) for modeling human judgements or opinions. In particular, such physical measures are assumed to be *observable indicators* of the dynamic process of rating which constitute the antecedents of the final rating outcome. In the DYFRAT approach these physical components are represented by: a) the motor activation involved in the process of rating and b) the overall time spent by the rater to provide h(er/is) final rating outcome. The first component is measured by means of tracking computer-mouse movements whereas the second component is measured in terms of response time - RT (recorded in ms). Finally, in our approach both the components are integrated into a common and comprehensive fuzzy model which allows to express the overall fuzziness of the combined temporal and motor dimensions of rating.

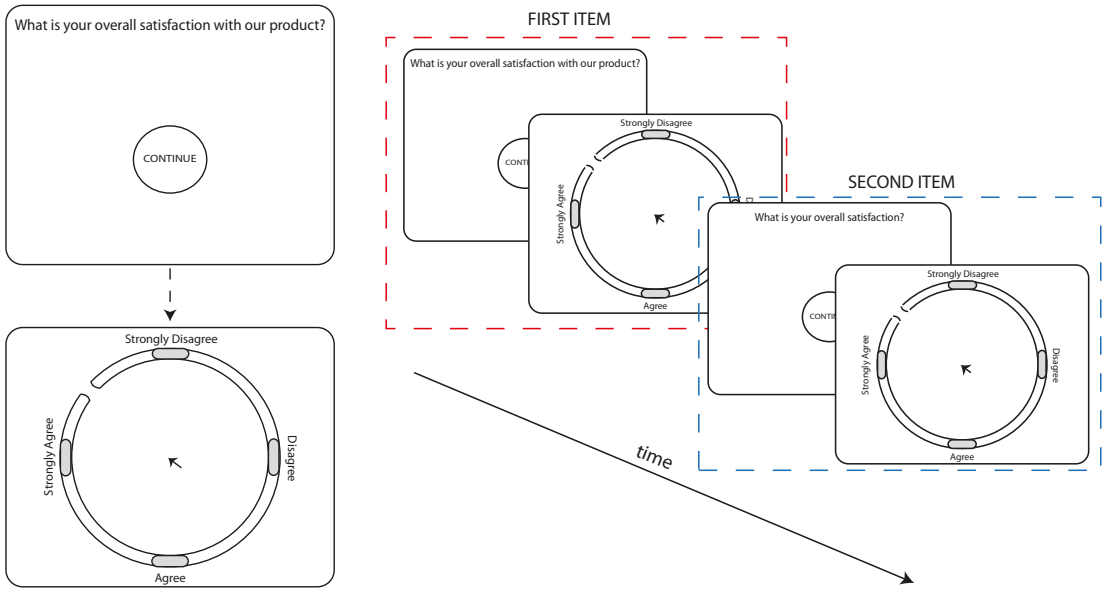
## 1.4 Dynamic Fuzzy Rating Tracker: methodology

DYFRAT consists of a *data-capturing procedure* which implements a MTM based computerized interface for collecting the motor and temporal components in the process of rating and a *data-modeling procedure* which provides a fuzzy model for the recorded information.

### 1.4.1 Data-capturing procedure

Consider a questionnaire with $J$ items. Each item is composed by its description (usually a question about a topic) and a rating scale with $K$ distinct levels (e.g., $K = 5$ for a five-point Likert scale). In the DYFRAT approach, a questionnaire is represented according to a customized graphical interface. In particular, for each item $j$ in the questionnaire, DYFRAT data-capturing interface is based on two consecutive frames (see figures 1.2a and 1.2b). After a participant clicks a start button, a window with the item description appears on the screen (*first frame*). A calibrated pause separates this first frame from the second frame. At the end of the pause, a pseudo-circular scale with $K$ levels appears while the mouse cursor is allocated to the center of the screen (*second frame*). A participant can give h(er/is) response by mouse-clicking the chosen level of the scale (the selected anchor point). Meanwhile, the DYFRAT system records the streaming of the x-y coordinates of the computer mouse (at a given sampling rate) as well as the total response time as the difference between the time at the mouse-clicking on the selected anchor point and the onset time of the presentation of the pseudo-circular scale on the screen. Note that the two frame-phases allow to separate information about the *reading process* (involved during the presentation of the item description) and the *response process* (involved during the presentation of the rating scale). This distinction ensures that mouse movements and response times can be related to distinct cognitive components. The idea behind the usage of a pseudo-circular scale in place of a traditional linear-type scale or arc-type scale (Freeman and Ambady, 2010) is justified by the fact that a pseudo-circular

configuration yields the largest degrees of freedom for mouse-movement recordings (Johnson et al., 2012). However, unlike genuine circular scales, a pseudo-circular representation still preserves the ordinal property between consecutive anchor points in the scale as it shows a visible breakpoint between the first level and the last level of the rating scale. The proposed configuration can be considered as a standard linear scale which has been bent in order to take into account the properties of both standard linear scales as well as circular ones.




(A) *General diagram of the proposed scale*

(B) *Temporal diagram for two or more questionnaire items*

FIGURE 1.2: DYFRAT: graphical schematization of the interface for data-capturing procedure

### 1.4.2 Data representation

The collected data are represented by an $I \times J$ array $\mathbf{D}$, that is to say, $I$ observations (raters) containing $J$ structured measures (rater's measures). The entry $d_{ij}$ of $\mathbf{D}$ denotes a *generalized rating measure* defined as:

$$d_{ij} = \langle z_{ij}, t_{ij}, \mathbf{p}_{ij} \rangle, \qquad i = 1, \ldots, I; j = 1, \ldots, J$$

with $z_{ij} \in \mathbb{N}$, $t_{ij} \in \mathbb{R}^+$, and $\mathbf{p}_{ij} = (\mathbf{x}_{ij}, \mathbf{y}_{ij}) \in \mathbb{R}^{N_{ij} \times 2}$ with $N_{ij} \in \mathbb{N}$. In particular, $z_{ij}$ represents the *final ordinal response* provided by rater $i$ to item $j$, $t_{ij}$ indicates the corresponding *response time*, whereas $\mathbf{x}_{ij}$ and $\mathbf{y}_{ij}$

are $N_{ij} \times 1$ vectors of Cartesian coordinates associated to the streaming of the x-y coordinates of the computer mouse movements (movements path of length $N_{ij}$) .
The main idea is to construct a new $I \times J$ array $\widetilde{\mathbf{D}}$, called the fuzzy data array of $\mathbf{D}$, by transforming each generalized rating measure $d_{ij}$ into a corresponding fuzzy set $\widetilde{d}_{ij}$ that expresses the overall fuzziness of the rating score associated with the $(i, j)$ response. Therefore, according to our rationale, we consider mouse-movements and response times as two particular sources of uncertainty, namely *spatial uncertainty* and *temporal uncertainty*. By contrast, the ordinal crisp response is understood as the final output of the dynamic decision process. A two-step procedure is used to derive the final fuzzy representation $\widetilde{d}_{ij}$ from $d_{ij}$. In the first step, we model (in an independent fashion) the spatial uncertainty and the temporal uncertainty. In the second step, we provide the final fuzzy model representation, $\widetilde{d}_{ij}$, by integrating these two sources of uncertainty. In particular, the rater's final response is captured by the core $core(\tilde{d}_{ij})$ of the fuzzy set $\tilde{d}_{ij}$, the spatial uncertainty is described by the support $supp(\tilde{d}_{ij})$, whereas the temporal uncertainty is modelled by the membership function $\mu_{\tilde{d}_{ij}}$.[1]

### 1.4.3 Data modeling

In what follows we describe the two-step procedure to derive the final fuzzy set $\widetilde{d}_{ij}$ according to the DYFRAT framework.

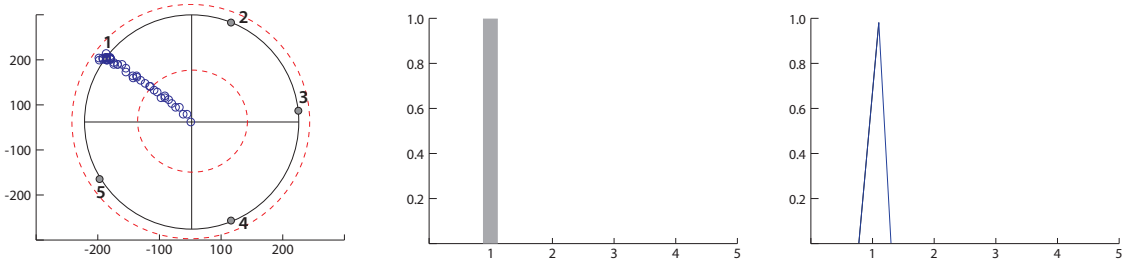
*First step*. In the first stage of the fuzzy modeling procedure, a fuzzy set $\widetilde{p}_{ij}$ representing the spatial uncertainty is constructed from the generalized rating measure $d_{ij}$. In particular, to derive $\widetilde{p}_{ij}$ we first remove eventual imprecision due to hand motor controls and/or computer mouse adjustments. To this end, the x-y coordinates in $\mathbf{p}_{ij}$ that are located near to the starting point (the center of the scale), the points that are recorded in the

[1] We recall that, given a collection $X$ of elements, the *support* of the fuzzy set $\tilde{d}$ on $X$ is the collection $supp(\tilde{d}) = \{x \in X \mid \mu_{\tilde{d}}(x) > 0\}$ whereas the *core* is the subset $core(\tilde{d}) = \{x \in X \mid \mu_{\tilde{d}}(x) = max_{z \in X}\ \mu_{\tilde{d}}(z)\}$.

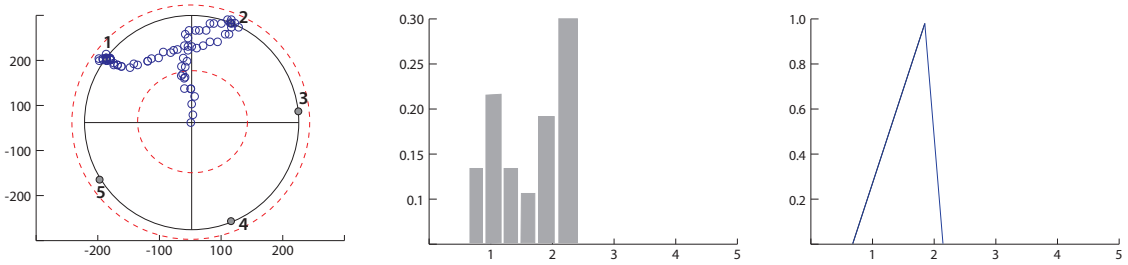
breakpoint area, and those ones that are beyond the border of the pseudo-circular scale, are all removed by applying a predefined filter which defines the area for acceptable x-y coordinates. The refined $\mathbf{p}_{ij}$ is next transformed into a vector $\mathbf{m}^c_{ij}$ of angles (expressed in radians) by using the well-known $atan_2$ function, namely $\mathbf{m}^c_{ij} = atan_2(\mathbf{p}_{ij})$.[2] The fuzzy set $\widetilde{p}_{ij}$ is constructed from the histogram of the radian measures collected in $\mathbf{m}^c_{ij}$. We call $\widetilde{p}_{ij}$ the spatial fuzzy set of $d_{ij}$. There are several procedures that can be adopted to derive fuzzy sets from data histograms (Medasani, Kim, and Krishnapuram, 1998). In the DYFRAT approach we adopted a heuristic procedure based on the *particle swarm optimization* (PSO) algorithm (Poli, Kennedy, and Blackwell, 2007). The PSO algorithm looks for the best fuzzy set that maximizes the total entropy with respect to the data histogram (Nieradka and Butkiewicz, 2007; Cheng and Chen, 1997; Li and Li, 2008). In particular, for the maximization algorithm we used the well known total fuzzy entropy measure proposed by De Luca and Termini (De Luca and Termini, 1972). Several convex as well as non-convex fuzzy sets can be used for representing a histogram (e.g., triangular, trapezoidal, gaussian. See: Ross, 2009; Calcagnì, Lombardi, and Pascali, 2013). However, for the sake of simplicity, in this contribution we opted for the simplest triangular format which is associated to the well-known LR representation (Dubois et al., 1988). In this respect, the triangular membership function may also be useful when one wants to analyze these variables on the basis of some widely used fuzzy statistical techniques (e.g., Taheri, 2003; Coppi, Gil, and Kiers, 2006). Some examples of fuzzy sets derived from the radian histograms are shown in figure 1.3. In particular, figure 1.3a shows a pattern of movements characterized by a modest spatial uncertainty which corresponds to a fuzzy set with a narrow support. By contrast, figure 1.3b shows a pattern of movements in which the spatial

[2] The $atan2(y, x)$ function is the arc tangent of the two variables $x$ and $y$ and uses the signs of both arguments in order to compute the quadrant of the result which lies in $(-\pi, \pi)$.

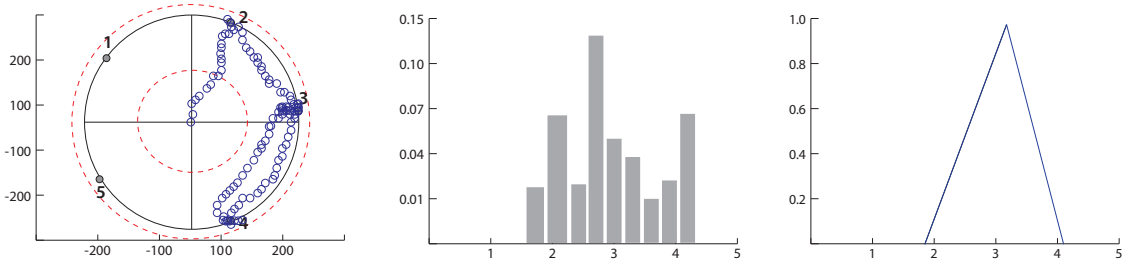
uncertainty is related to the rater's choice between two possible alternatives. Note that in this second configuration the associated fuzzy set has now a wider support and its core has shifted toward the right side. Finally, figure 1.3c presents an interesting pattern in which the spatial uncertainty is related to the choice among three distinct options. This last configuration shows the largest support for the derived fuzzy set.



(A) *Spatial uncertainty related to a simple direct response (final response = 1)*



(B) *Spatial uncertainty related to the choice between two alternatives (final response = 1)*



(C) *Spatial uncertainty related to the choice among three alternatives (final response = 3)*

FIGURE 1.3: Three types of empirical mouse movements with the correspondent histograms and fuzzy sets. Note that, the continuous black circle represents the original scale, the dashed red ones represents the filters whereas the blue points indicate the recorded mouse movements. The ordinal numbers codify the anchor points of the scale (e.g., Strongly Disagree = 1, Disagree = 2, Neither = 3, Agree = 4, Strongly Agree = 5) that are simply juxtaposed on the scale of radians for the sake of exposition.

*Second step*. In the second modeling step, the final fuzzy representation, is obtained by transforming the spatial fuzzy set $\widetilde{p}_{ij}$ into a new fuzzy set $\widetilde{d}_{ij}$ which integrates the spatial component with the temporal information of the rating process. In particular, the membership function of $\widetilde{d}_{ij}$ is defined

according to the following conditional equations:

$$\mu_{\tilde{d}_{ij}}(x) = \begin{cases} 0.5 - (2^{2\gamma_{ij}} - 1) \cdot [0.5 - \mu_{\tilde{p}_{ij}}(x)]^{2\gamma_{ij}}, & \widehat{F}(t_{ij}) > \widehat{F}(\bar{t}_j) \quad \text{and} \quad 0 \leq \mu_{\tilde{p}_{ij}}(x) \leq 0.5 \\ 1 - (2^{0.5\omega_{ij}}) \cdot [1 - \mu_{\tilde{p}_{ij}}(x)]^{0.5\omega_{ij}}, & \widehat{F}(t_{ij}) > \widehat{F}(\bar{t}_j) \quad \text{and} \quad 0.5 < \mu_{\tilde{p}_{ij}}(x) \leq 1 \\ \left[\mu_{\tilde{p}_{ij}}(x)\right]^{\nu_{ij}}, & \widehat{F}(t_{ij}) < \widehat{F}(\bar{t}_j) \quad \text{and} \quad 0 \leq \mu_{\tilde{p}_{ij}}(x) \leq 1 \\ \mu_{\tilde{p}_{ij}}(x), & \widehat{F}(t_{ij}) = \widehat{F}(\bar{t}_j) \quad \text{and} \quad 0 \leq \mu_{\tilde{p}_{ij}}(x) \leq 1 \end{cases} \tag{1.1}$$

with

$$\gamma_{ij} = \lambda_{ij}\widehat{F}(t_{ij}), \quad \omega_{ij} = \frac{1}{\gamma_{ij}}, \quad \text{and} \quad \nu_{ij} = \frac{1}{2\widehat{F}(t_{ij})}$$

and where $\widehat{F}$ denotes the empirical cumulative distribution function of the response times sample $\mathbf{t}_j = (t_{1j}, t_{2j}, \ldots, t_{Ij})$ associated to item $j$ whereas $\bar{t}_j$ indicates the sample mean of $\mathbf{t}_j$. The first and the second lines of formula 1.1 act as an *expansion* modifier which increases the overall uncertainty represented in the original spatial fuzzy set $\widetilde{p}_{ij}$. In particular, if the recorded time $t_{ij}$ for rater $i$ to item $j$ is larger than the corresponding average time $\bar{t}_j$, then $\widetilde{p}_{ij}$ is nonlinearly expanded according to a model parameter $\gamma_{ij}$ which, in turn, depends on the cumulative density value of $t_{ij}$ and the shape parameter $\lambda_{ij}$[3]. By contrast, the third line of formula 1.1 acts as a *concentration* modifier which decreases the overall uncertainty represented in $\widetilde{p}_{ij}$. In particular, if the recorded time $t_{ij}$ for rater $i$ to item $j$ is smaller than the corresponding average sample time $\bar{t}_j$, then $\widetilde{p}_{ij}$ is nonlinearly concentrated according to a shape parameter $\nu_{ij}$ which is inversely related to the cumulative density value of $t_{ij}$. Finally, the fourth line of formula 1.1 represents the fuzzy set when the observed response time $t_{ij}$ is perfectly equivalent to the average response time for item $j$. In this last case the new fuzzy set $\widetilde{d}_{ij}$ simply boils down to the original spatial fuzzy set $\widetilde{p}_{ij}$. Note that, that the third line of formula 1.1 corresponds to the well-known linguistic hedge called *concentration* which allows to reduce the fuzziness of the set. Similarly, the first and second lines of formula 1.1 is inspired by the

[3] $\lambda_{ij}$ is a coefficient which maximizes the overall fuzziness of the set. Given the fuzzy set of movements $\widetilde{p}_{ij}$, the best value for $\lambda_{ij}$ is obtained by adopting an iterative optimization algorithm which maximizes the Kaufmann index (Kaufmann and Swanson, 1975). Because the objective function is bounded above, the algorithm is always able to yield the maximum of the function. However, for basic triangular representations $\lambda_{ij}$ can be simply chosen to be $\geq 4$.

Zadeh's linguist hedge called *intensification* (Huynh, Ho, and Nakamori, 2002) which allows to intensify (to expand) the fuzziness of the set. In our approach the expansion was created ad-hoc in order to obtain an increasing level of fuzziness for the set. More precisely, if $\mu_{\widetilde{p}_{ij}}(x) \leq 0.5$ we expanded the base of the set, otherwise we reduced the peak of the set. Figure 1.4 shows a graphical summary of this expansion/concentration transformation.

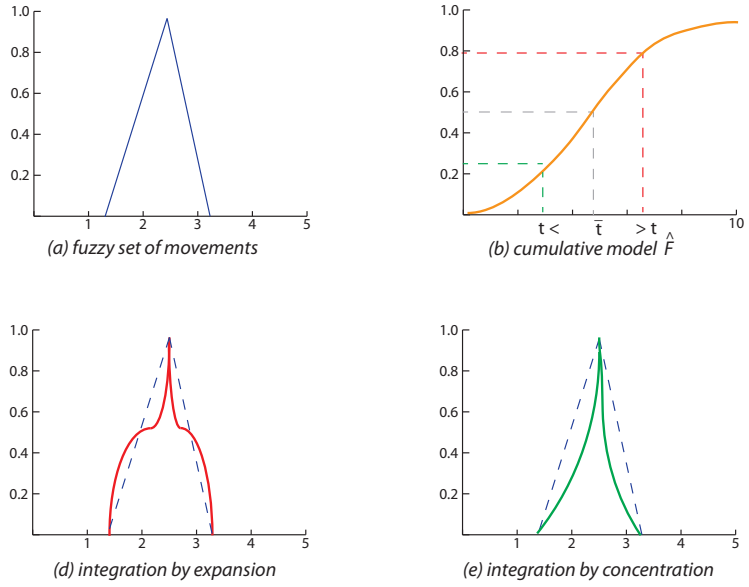



FIGURE 1.4: Graphical schematization for the integration process

However, in the DYFRAT context we stress that the equations used for the integration between movement patterns and response times are only used according to a descriptive fashion and, therefore, they cannot be conceived as linguistic hedges in a strict fuzzy logic sense. Moreover, we preferred to develop a novel fuzzy modifier according to a statistically oriented perspective (e.g., using the cumulative distribution function for the response times and the observed average response time for the item) because we were interested in providing a rational procedure to transform a fuzzy set according to observed distributional data (empirical distribution of times and empirical distribution of movements).

In sum, the overall idea underlying the definition of the novel modifier was based on current theories about decision making (e.g., Greene and

Haidt, 2002). In particular, according to these theories intuitive judgements or evaluations occur quickly, effortlessly, and almost automatically, such that the final responses (but not necessarily the underlying processes) are accessible to consciousness. By contrast, more elaborated or conflicting reasonings occur more slowly, require some additional efforts, and presumably involve some more steps that are directly accessible to consciousness. The latter type of responses are usually characterized by a much larger level of uncertainty. With our model we tried to capture this psychological intuition according to a purely descriptive representation.

### 1.4.4 Summary measures

Fuzzy summary measures can play a relevant role in highlighting important properties of the final fuzzy set $\widetilde{d}_{ij}$. In the DYFRAT framework we implemented the following basic and well known fuzzy set measures:

- *Kaufmann index* (Kaufmann and Swanson, 1975):

$$K(\widetilde{d}_{ij}) = \frac{2}{\text{card}(\widetilde{d}_{ij})} \cdot \sum_x \mid \mu_{\widetilde{d}_{ij}}(x) - \delta(x) \mid \tag{1.2}$$

$$\text{with} \quad \delta(x) = \begin{cases} 1, & \text{if } \mu_{\widetilde{d}_{ij}}(x) \geq 0.5 \\ 0, & \text{if } \mu_{\widetilde{d}_{ij}}(x) < 0.5 \end{cases}$$

  and where card(.) is the cardinality of the final fuzzy set $\widetilde{d}_{ij}$.

- *Fuzzy entropy* De Luca and Termini, 1972:

$$H(\widetilde{d}_{ij}) = -\sum_x [\mu_{\widetilde{d}_{ij}}(x) \log(\mu_{\widetilde{d}_{ij}}(x))] - [(1-\mu_{\widetilde{d}_{ij}}(x)) \log(1-\mu_{\widetilde{d}_{ij}}(x))] \tag{1.3}$$

- COG based *Fuzzy centroid* Ross, 2009:

$$CR(\widetilde{d}_{ij}) = \left(\sum_x x \cdot \mu_{\widetilde{d}_{ij}}(x)\right) \cdot \left(\sum_x \mu_{\widetilde{d}_{ij}}(x)\right)^{-1} \tag{1.4}$$

- *Total spread (or length of the fuzzy set support)*:

$$TS(\widetilde{d}_{ij}) = max(supp(\widetilde{d}_{ij})) - min(supp(\widetilde{d}_{ij})) \tag{1.5}$$

In addition, we also considered a new simple fuzzy measure, called the *intensification index*, derived from the fuzzy entropy measure. In particular, the intensification index is defined as follows:

$$H_R(\widetilde{d}_{ij}, \widetilde{p}_{ij}) = \frac{H(\widetilde{d}_{ij}) - H(\widetilde{p}_{ij})}{H(\widetilde{p}_{ij})} \tag{1.6}$$

where $H(\widetilde{p}_{ij})$ and $H(\widetilde{d}_{ij})$ indicate the entropies associated to the spatial and final fuzzy sets, respectively. Values of $H_R(\widetilde{d}_{ij}, \widetilde{p}_{ij}) < 0$ indicate the quantity of information which is subtracted by concentration from the spatial fuzzy set $\widetilde{p}_{ij}$, whereas values of $H_R(\widetilde{d}_{ij}, \widetilde{p}_{ij}) > 0$ indicate the quantity of information which is added by expansion on $\widetilde{p}_{ij}$.
One important comment is in order concerning the defuzzification measures adopted for the final fuzzy set. In general, several measures can be selected to perform the defuzzification of a fuzzy set (Roychowdhury and Pedrycz, 2001). In this first implementation of the DYFRAT system we opted for a COG based index because of its simplicity and high flexibility. In particular, the COG centroid is a measure which fully takes into account the integration between movements and times. More specifically, unlike other defuzzification measures (e.g., first of maximum FoM, last of maximum LoM, mean of maxima MEoM, etc.), the COG index also considers the weighted information provided by the membership function of the final fuzzy set.
In sum, these measures can be used, among other things, for quantifying some peculiar information stored in the final fuzzy sets (e.g., the amount of ambiguity, fuzziness, uncertainty, average information), for detecting eventual outliers in the data (i.e., subjects that show anomalous spatial temporal information), as well as for running further statistical analysis. In particular, the total spread can be used for detecting those individuals

who use the computer-mouse in an improper way (outliers for mouse-movements). Similarly, the intensification index can be used for detecting anomalous subjects with lower or greater response times (outliers for response times). We will provide some examples of indices applications in the sixth section of the manuscript.

## 1.5 Dynamic Fuzzy Rating Tracker: implementation

*Data-capturing procedure*. The first application consists in an executable stand-alone package available for Windows, OSX and Unix systems developed in Processing 2.0 (http://processing.org). All the main features of the DYFRAT graphical interface can be modified by the user (e.g., temporal delay between-items or within-items, labels positions, font type and text size, scale diameter, scale stroke, breakpoint width, mouse-movements sample-rate). In particular, label positions can be set to either a fixed or random configuration. In the latter case, the position of the first label is located at random on the circular scale and all the other labels are reconfigured accordingly to reconstruct the linear order. Moreover, the user can also modify the questionnaire by choosing the number of items and the levels of the scale as well as the textual information describing the scale levels. The user can manage the application by modifying specific textual files (placed in the application main folder) that contain all the application parameters. When the application is launched, the files are automatically loaded. Figure 1.5-a shows an example of the final DYFRAT graphical interface. For each combination of rater $i$ and item $j$, the application provides two output files (textual format). The first output contains information related to the rater's responses (mouse movements in cartesian coordinates, degrees, ordinal crisp values for the final response, response time, on-set and off-set values). The second output is instead a pre-formatted textual file containing information that is used as input for the data-modeling application.

*Data-modeling procedure*. The second application consists of a GUI-based system developed in Matlab for Windows, OSX and Unix systems that implements the modeling steps of the DYFRAT methodology (see Figure 1.5-b). All the features involved in the analysis (e.g., type of filters, histogram of movements, PSO parameters) can be set by the user. Moreover, different types of analysis (e.g., subject-by-subject, item-by-item, global) with different characteristics (temporized or static analysis) can be selected. Finally, the application provides a single output containing the main results of the analysis. These are organized by means of two-way (two-dimensional matrices) as well as three-way array structures (three dimensional matrices).

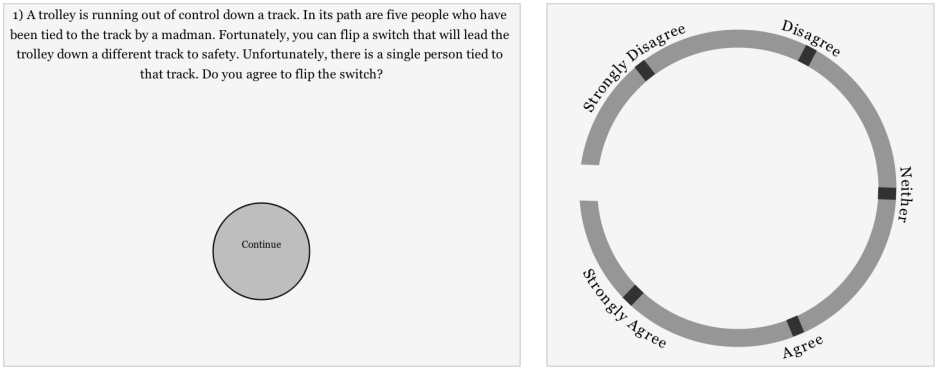
## 1.6 Illustrative examples

By way of illustration we consider three simple applications using the DYFRAT methodology. The first example is about the evaluation of a well known cognitive problem in decision making. The second application considers data about rash driving behaviors among young people aged 18-26. Finally, the third application illustrates how one can perform an outlier detection analysis using the DYFRAT framework.

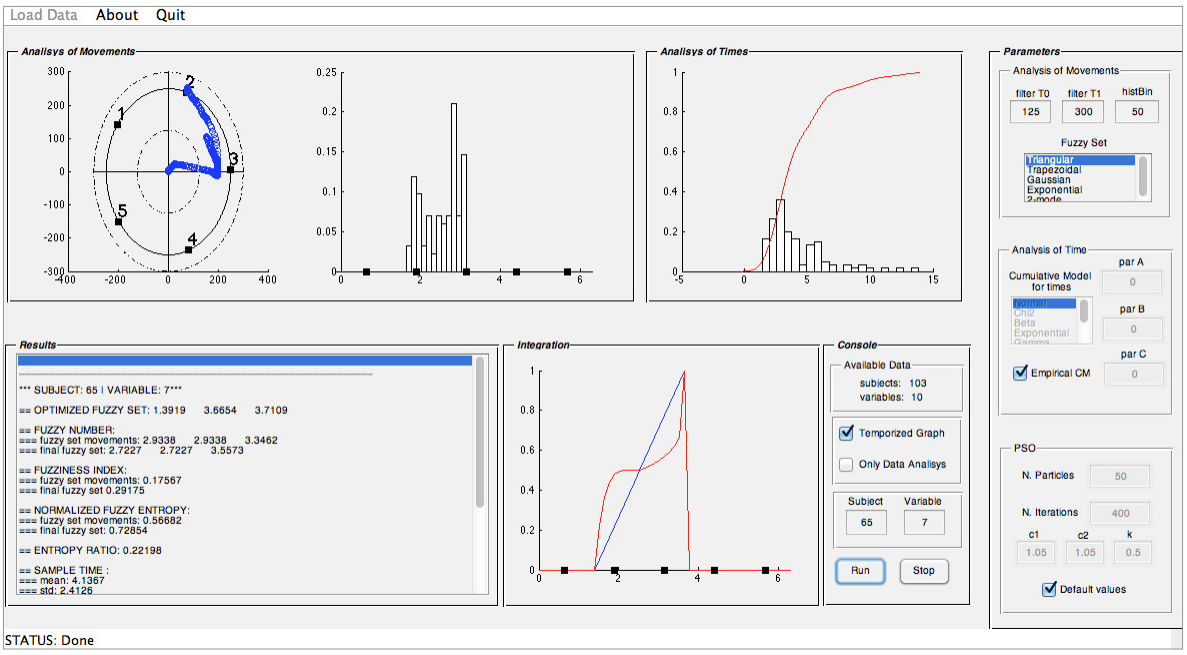
### 1.6.1 Rating responses and moral dilemma

*General context and motivation*. In cognitive decision making (Greene and Haidt, 2002; Haidt, 2001), moral judgements and dilemmas are relevant phenomena characterized by high levels of uncertainty in individuals responses. In this application, we used a moral dilemma based on the well-known *trolley scenario* (Greene and Haidt, 2002; Haidt, 2001; McGuire et al., 2009):

> A trolley is running out of control down a track. In its path are five people who have been tied to the track by a madman. Fortunately, you can flip a switch that will lead the trolley down a different track to safety. Unfortunately, there is a single person tied to that track. Do you agree to flip the switch?




(A) *Interface for data-capturing developed in Processing*




(B) *Interface for data-modeling developed in Matlab*

FIGURE 1.5: Screen-shots from DYFRAT implementation

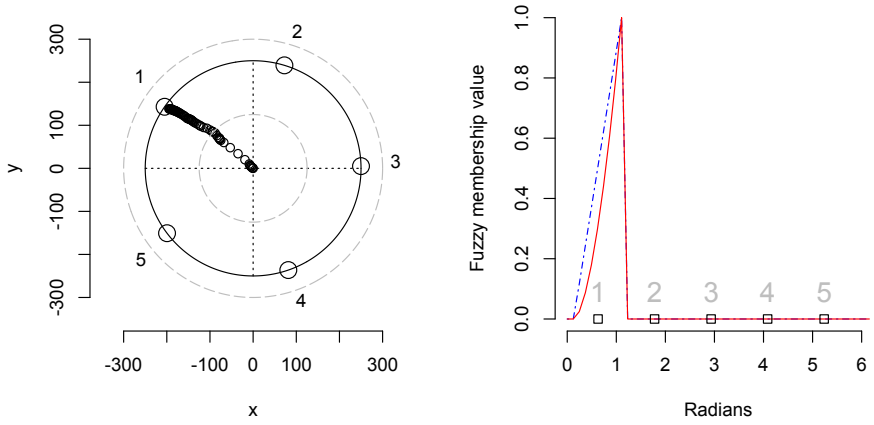
In general, individuals react to the dilemma by using their personal moral beliefs. The trolley scenario is an example of a setting engaging some levels of emotional processing which, in turn, might affect people's judgments to such an extent that some individuals may hesitate in providing the final response. Researchers working in decision making (e.g., Haidt, 2001; Bruner and Bruner, 2009) stress the fact that moral evaluations can be distinguished in two components: moral intuition and moral reasoning. Moral intuition occurs quickly, effortlessly, and automatically, such that

the final judgment (but not the underlying process) is accessible to consciousness, whereas moral reasoning occurs more slowly, requires some effort, and involves at least some steps that are accessible to consciousness. In this first example, we studied the relationship between response uncertainty, as measured by the Kaufmann index of the final fuzzy set, and moral judgement as represented by the centroid of the same fuzzy set. We expect that the individuals who show a very strong disagreement with the action described in the trolley scenario (moral intuition raters) will be characterized by very fast responses with low levels of uncertainty. By contrast, those who are characterized by a more moral thinking attitude (moral reasoning raters) will show less extreme responses with larger values of uncertainty. Because the trolley scenario may activate not necessarily conscious underlying processes in the rater, we believe that DYFRAT can represent an ideal methodology for testing this hypothesis.

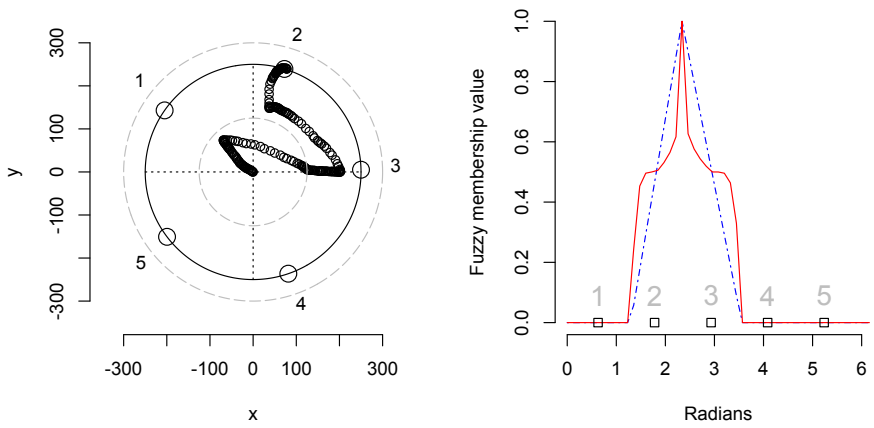
*Data-analysis and results*. The trolley dilemma was administered to a group of students ($I = 103$, 47 males, age 18-23 : $70.87\%$, age 24-27 : $19.42\%$, age 28-36 : $2.91\%$, age $\geq 37$ : $3.88\%$) from the University of Trento (Italy) and the responses were collected using the DYFRAT graphical interface. In particular, participants used a pseudo-circular scale with five response levels (strongly disag.=1, disag.=2, neither=3, agree=4, strongly agree=5). Figure 1.6 shows two empirical patterns of mouse movements with the final fuzzy sets. In particular, Figure 1.6a represents an empirical pattern with a low uncertainty/fuzziness, by contrast Figure 1.6b shows a pattern with a higher level of uncertainty. Table 1.1 reports some results for the two selected subjects.

| Subj. | $z$ | CR | $t$ | K | H | $H_R$ | TS |
|---|---|---|---|---|---|---|---|
| 21 | 1 | 0.9 | 2.62 (4.72) | 0.06 | 3.34 | -0.15 | 1.03 |
| 80 | 2 | 2.44 | 6.25 (4.72) | 0.30 | 11.53 | 0.21 | 3.61 |

TABLE 1.1: Example 1: DYFRAT results ($z$=discrete response, CR=fuzzy centroid, $t$=subject response time with sample time in parenthesis, K=Kaufmann index, H=entropy, $H_R$=intensification index, TS=total spread)




(A) *Subject 21*




(B) *Subject 80*

FIGURE 1.6: Example 1: Empirical patterns of mouse movements with the associated final fuzzy sets. Note that, the dotted grey circle represents the filters whereas the ordinal numbers on the circles and those ones on the radians scale represent the anchor points (Strongly Disagree = 1, Disagree = 2, Neither = 3, Agree = 4, Strongly Agree = 5)

The scatterplot between response uncertainty and moral judgement evaluation is presented in Figure 1.7. The result of the graphical analysis shows that the raters with fuzzy centroid values closer to the extreme anchor point "strongly disagree" are characterized by very low levels of response uncertainty (for these individuals the Kaufmann index is very low). By contrast, raters with fuzzy centroid values closer to less extreme or intermediate anchor points (disagree or neither) showed much larger levels of

uncertainty in the rating process. In particular, it is interesting to note a

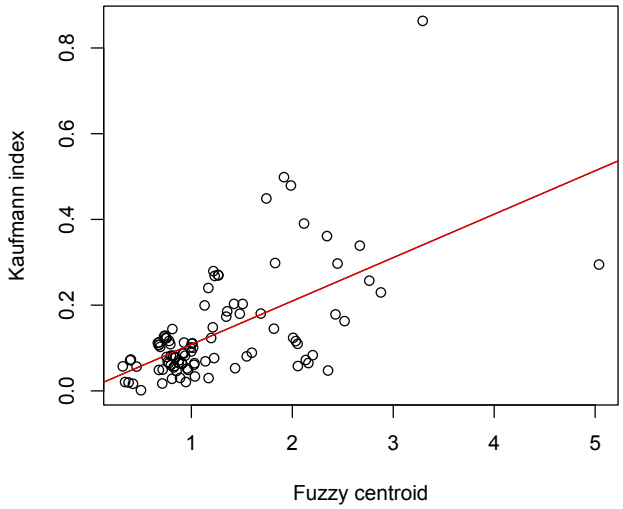



FIGURE 1.7: Example 1: Scatter plot between Kaufmann index and fuzzy centroid

positive linear trend between these two variables (the higher the value of the rating judgement, the larger the level of the observed response uncertainty). The linear model fitted on data showed a good fit ($R^2 = 0.6$) and a statistically significant relation between the two variables ($\beta_{\text{CR}} = 0.39$, $p < .01$). This result is in line with the theoretical expectation that moral intuition raters are characterized by more extreme, quick, effortless, and automatic final responses. In sum, this application shows how DYFRAT can be considered as an efficient and elegant procedure to represent the underlying mechanisms involved in the cognitive process of rating in moral dilemmas.

### 1.6.2 Self-report behaviors in reckless driving

*General context and motivation*. In this second application we studied the effects of gender and driving experience on self-report measures about reckless driving in a group of young adults. High rates of traffic fatalities among young people continue to be a very serious, worldwide problem (e.g., Arnett, Offer, and Fine, 1997). Nowadays, there are several evidences that show how young drivers are generally more likely to perform

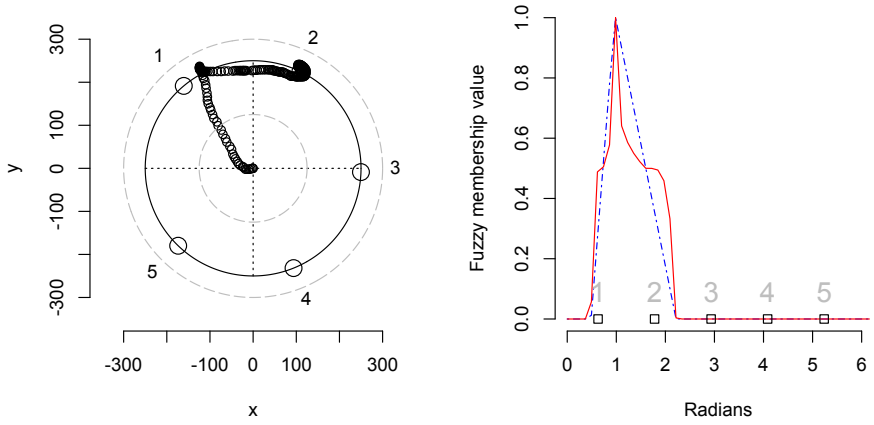
risky behaviors or neglect precautions while driving than more experienced drivers (Arnett, Offer, and Fine, 1997; Jonah, 1986). In particular, several studies have shown that unexperienced young drivers ability to perceive risk accurately is generally low (Finn and Bragg, 1986; Glendon et al., 1996). Moreover, young men seem to consider reckless driving less serious than do young women (DeJoy, 1992) and, to some extent, this tendency seems not to be necessarily rationally (or consciously) based (Deery, 2000). Because in this sensitive context, self-report ratings can be influenced by implicit aspects, we used the DYFRAT interface to track and collect real-time behavioral data occurring during the rating process.

*Data analysis and results*. A six-item questionnaire was adapted from a previous reckless driving scale (Taubman-Ben-Ari, Mikulincer, and Iram, 2004) and administered to a group of young drivers ($I = 60$, 38 males, age 18-23 : $43.33\%$, age 24-28 : $28.33\%$, age $\geq 29$ : $28.33\%$) from the Trentino region (North-East Italy). The only criteria for inclusion in the study were possession of a driving license and at least six months of driving experience. Table 1.2 reports the item descriptions. Participants were asked to read each item carefully and report how often they used to drive according to the described way. Data were collected using the DYFRAT graphical interface and ratings were made on a 5-point scale, ranging from 1 (*never*) to 5 (*very often*). In Figures 1.8 and 1.9 we illustrate four empirical patterns (two females and two males) on the second item only. Table 1.3 shows the DYFRAT results for these selected cases.

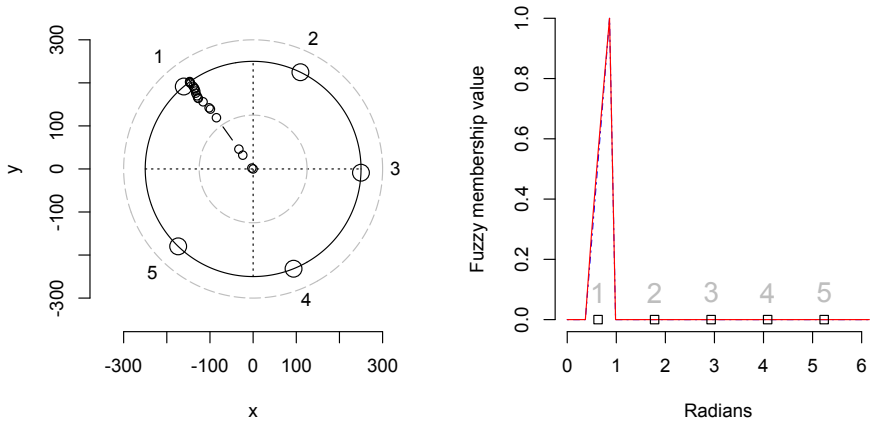
| Item | Description |
|---|---|
| 1 | Parking in a non-parking zone |
| 2 | Not stopping in a stop sign |
| 3 | Overtaking another vehicle on a continuous white line (no pass zone) |
| 4 | Not keeping the right distance from the vehicle in front of me |
| 5 | Driving under the influence of alcohol |
| 6 | Turning in high speed |

TABLE 1.2: Example 2: Reckless driving questionnaire

The dependent variables of the study were the fuzzy summary measures




(A) *Subject 5*

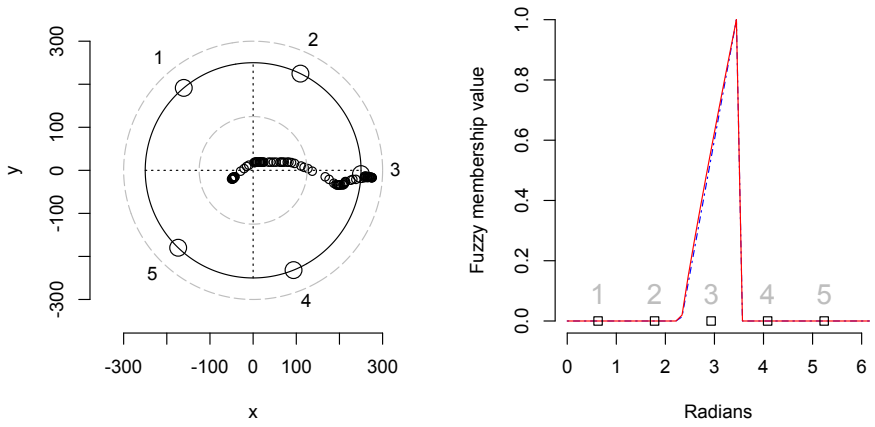



(B) *Subject 37*

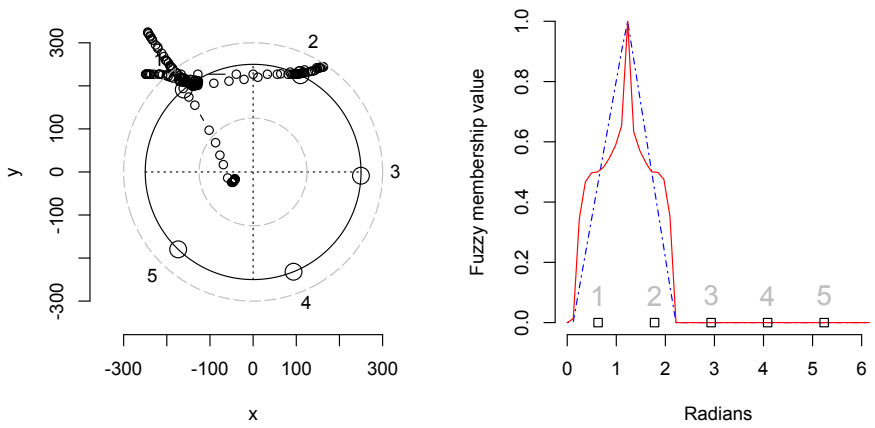
FIGURE 1.8: Example 2: Empirical patterns of mouse movements with the associated final fuzzy sets on the second variable (female only). Note that, the dotted grey circle represents the filters whereas the ordinal numbers on the circles and those ones on the radians scale represent the anchor points (Strongly Disagree = 1, Disagree = 2, Neither = 3, Agree = 4, Strongly Agree = 5)

| Subj. | $z$ | CR | $t$ | K | H | $H_R$ | TS |
|---|---|---|---|---|---|---|---|
| 5 (F) | 2 | 1.31 | 6.07 (3.82) | 0.21 | 8.42 | 0.19 | 3.25 |
| 37 (F) | 1 | 0.8 | 3.26 (3.82) | 0 | 1.82 | 0 | 0.60 |
| 41 (M) | 3 | 3.16 | 3.26 (3.82) | 0.08 | 4.56 | 0 | 1.16 |
| 46 (M) | 1 | 1.20 | 4.06 (3.82) | 0.25 | 10.25 | 0.20 | 2.15 |

TABLE 1.3: Example 2: DYFRAT results for the second variable only (M=male, F=female, $z$=discrete response, CR=fuzzy centroid, $t$=subject response time with sample time in parenthesis, K=Kaufmann index, H=entropy, $H_R$=intensification index, TS=total spread




(A) *Subject 41*




(B) *Subject 46*

FIGURE 1.9: Example 2: Empirical patterns of mouse movements with the associated final fuzzy sets on the second variable (male only). Note that, the dotted grey circle represents the filters whereas the ordinal numbers on the circles and those ones on the radians scale represent the anchor points (Strongly Disagree = 1, Disagree = 2, Neither = 3, Agree = 4, Strongly Agree = 5)

(fuzzy centroid, total spread, fuzzy entropy, Kaufmann index, intensification index) and the crisp rating response, whereas the independent variables were the factors gender and driving experience (at two levels: $< 3$; $\geq 3$ years). A *t*-test for independent samples was separately performed for each item in the questionnaire. The results of the analysis are reported in Table 1.4.

As expected, the observed differences were in line with what documented

| Item | $z$ | CR | K | H | $H_R$ | TS |
|---|---|---|---|---|---|---|
| 1 | $-1.84^{*}_{g}$, $-0.19_{e}$ | $-2.25^{**}_{g}$, $-0.57_{e}$ | $0.15_{g}$, $-0.47_{e}$ | $0.32_{g}$,$-0.54_{e}$ | $-0.02_{g}$,$-1.96_{e}$ | $0.98_{g}$,$-0.18_{e}$ |
| 2 | $-0.92_{g}$, $-0.36_{e}$ | $-1.53_{g}$, $-0.84_{e}$ | $-2.57^{**}_{g}$, $0.10_{e}$ | $-2.67^{**}_{g}$,$0.03_{e}$ | $-1.92^{*}_{g}$,$-0.64_{e}$ | $-2.29^{**}_{g}$,$0.24_{e}$ |
| 3 | $-1.87^{*}_{g}$, $1.04_{e}$ | $-1.45_{g}$, $0.18_{e}$ | $-1.99^{**}_{g}$, $1.39_{e}$ | $2.06^{**}_{g}$,$1.23_{e}$ | $-0.84_{g}$,$-0.31_{e}$ | $2.67^{**}_{g}$,$1.05_{e}$ |
| 4 | $-0.23_{g}$, $-0.07_{e}$ | $-0.06_{g}$, $0.60_{e}$ | $2.21^{**}_{g}$, $0.23_{e}$ | $2.26^{**}_{g}$,$0.23_{e}$ | $0.50_{g}$,$-0.57_{e}$ | $2.08^{**}_{g}$,$0.29_{e}$ |
| 5 | $-1.65^{*}_{g}$, $1.00_{e}$ | $-1.10_{g}$, $0.64^{**}_{e}$ | $-0.50_{g}$, $-0.59^{*}_{e}$ | $-0.42_{g}$,$-0.53_{e}$ | $-0.98_{g}$,$-1.11^{**}_{e}$ | $-0.14_{g}$,$-0.10_{e}$ |
| 6 | $-2.61^{**}_{g}$, $1.59_{e}$ | $-2.55^{**}_{e}$, $2.25_{g}$ | $-1.36_{g}$, $-1.82_{e}$ | $-1.49_{g}$,$-1.62_{e}$ | $-0.71_{g}$,$-2.32_{e}$ | $-1.54_{g}$,$-1.04_{e}$ |

TABLE 1.4: $t$-test for independent samples ($t_{AB}$) observed statistic values for the crisp rating variable (z) and the fuzzy summary variables (CR: fuzzy centroid, H: entropy, $H_R$: intensification index, TS: total spread, K: fuzziness). The subindex $g$ denotes a $t_{AB}$ statistic computed for the gender factor ($A$ : females, $B$ : males). The subindex $e$ denotes a $t_{AB}$ statistic computed for the experience driving factor ($A$ : not expert drivers, $B$ : expert drivers). If $t_{AB} < 0$ (resp. $t_{AB} > 0$), then the r.v. for group $A$ (resp. group $B$) is stochastically larger than the r.v. for group $B$ (resp. group $A$). Note that: $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$

in the reckless driving literature with male drivers reporting more frequently risky behaviors than female drivers. However, more interesting differences between the two groups emerged when the analysis was repeated using the fuzzy summary statistics as dependent variables. In particular, sensation seeking and more aggressive driving styles (items 3 and 6) were found to be more characteristic of young male drivers and less of young female drivers. In addition, the male group reported on a higher appraisal of driving as a challenge than the female group (item 5). In general, it seemed that young male drivers tended to disregard more potential negative outcomes in comparison to women. By contrast, young female drivers tended to perceive driving as more threatening in comparison to male. Finally, an illuminating difference was observed between not expert drivers (with less than 3 years of driving experience) and more expert drivers (with at least 3 years of driving experience). In particular, not expert drivers considered alcohol consumption (item 6) less serious and less likely to result in a dangerous source for potential harm than the more expert-drivers reported.

### 1.6.3 Outlier detection analysis

*General context and motivation*. In this last example we illustrate how the DYFRAT methodology can be used to perform outlier detection analysis in a dynamic rating setting. In general, in a data sample an outlier can be defined as a data point or data observation that appears to deviate markedly from other observations in the sample (Grubbs, 1969; Barnett and Lewis, 1994). Of course, this definition implicitly entails a sort of distance metric to measure single data-point deviation from the data sample. An alternative and more general definition for outliers is a subset of observations which appears to be inconsistent with the remainder of observations in the sample (Barnett and Lewis, 1994). In our context, outliers refer to anomalous measure values describing the temporal and spatial dynamics underlying the mechanism involved in the cognitive process of rating. In this respect, two particular fuzzy summary measures, the total spread and the

intensification index, seem to be appropriate statistics to detect eventual anomalies in the spatial and temporal components of the rating process as measured by computer-mouse movements and relative response time, respectively.

*Data analysis and results*. For the sake of simplicity, here we illustrate the outlier detection procedure using only the data associated to the fifth item of the reckless questionnaire described in the former application. We recall that this item described a situation where an individual drives under the influence of alcohol (see Table 1.2).

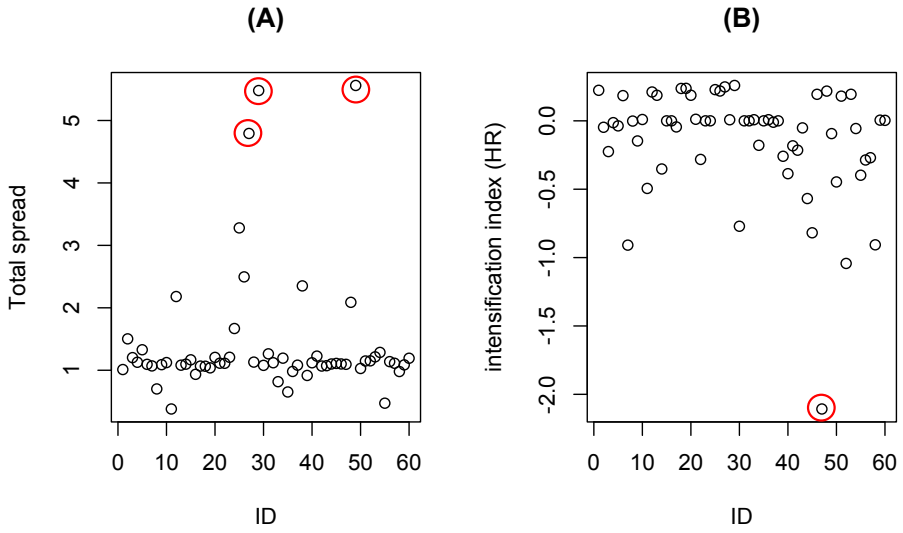



FIGURE 1.10: Example 3: Scatter plots for total spread and intensification index for outlier detection. Note that, red circles indicate outliers.

Figure 1.10A shows the total spread measure as a function of the subject identification number (ID) in the questionnaire. Note that there are three data points (outliers) which are clearly isolated and inconsistent with the main cluster of points. These outliers correspond to raters that have shown patterns of mouse movements that markedly deviate from the other spatial patterns in the sample. Figure 1.11 shows the corresponding empirical patterns of mouse movements. Figure 1.10B shows the intensification index $H_R$ as a function of subject ID. In this graphical representation there is only one rater ($i = 47$) who is characterized by a very fast response

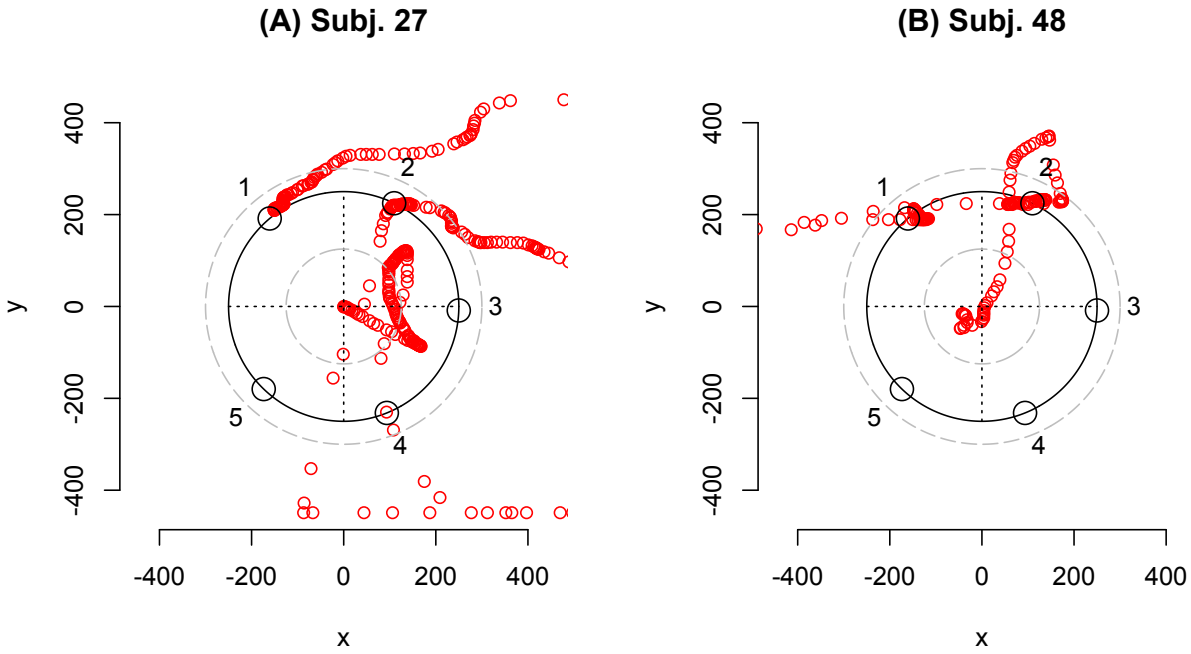



FIGURE 1.11: Example 3: Anomalous empirical patterns of mouse movements.

time relative to the response times variance for the group of raters. Finally, we remark that in some situations it can be useful to produce a scatter plot between the total spread statistic and the intensification index. This combined representation would be used to jointly detect eventual anomalies in the spatial and temporal components of the rating process.

## 1.7 Final remarks

In this paper we proposed a novel methodology (DYFRAT) for measuring the fuzziness of human rating situations from an original perspective. By considering human rating as a temporal and dynamic changing course of information in which the final rater's response is only the outcome of peculiar latent cognitive processes, we modelled fuzziness as the result of the integration between two important physical/biometric measures. We adopted the mouse tracking methodology for capturing the motor and

temporal components of the rating process. Such components were computed by the streaming of x-y computer-mouse coordinates and the overall response times, respectively. In order to provide a final model representation for the rating data, we integrated such information using a fuzzy modeling paradigm. In particular, the motor component was represented by an appropriate fuzzy set whereas the temporal component was included by modifying the shape of the spatial fuzzy set. The final fuzzy representation included all the information available from the response process (i.e., the final response, the temporal and spatial information, the fuzziness of the response, etc). To better illustrate the DYFRAT features, we also described two real applications from decision making and risk assessment contexts. The results suggested how DYFRAT can measure important features of dynamic decision process. We further showed as DYFRAT could be used to perform an outlier detection analysis.

### 1.7.1 Limitations

As with other fuzzy rating approaches (e.g., FCS or FRS), also DYFRAT involves simplifying assumptions that may result in some concerns about the specific procedure adopted for the fuzzy variables' construction. For example, the choice to represent the shape of the spatial fuzzy set $\widetilde{p}_{ij}$ according to a triangular membership function may appear unjustified for some empirical applications. However, a straightforward but more flexible generalization could be obtained by using, for example, multimodal representations for the spatial fuzzy set (e.g., Calcagnì, Lombardi, and Pascali, 2013). Nonetheless, it is not difficult to see that also this more sophisticated representation would still require an analyst to set up some ad-hoc options in the construction of the membership functions. A future venue of research should be dedicated to the exploration and analysis of alternative representations for the spatial fuzzy set to minimize the need for ad-hoc representational assumptions. Another possible extension of the DYFRAT approach would also consider alternative perspectives for the approximation of the dynamics of rating on the basis of more qualitative

representations (Chung and Schwartz, 1995). Finally, at a more methodological level a better validation of the advantages of our proposal compared with other existing methods for fuzzy ratings (e.g., FCS and FRS) should be tested in future works.

### 1.7.2 Conclusions

In sum, unlike other fuzzy scales, DYFRAT allows to express the fuzziness of the human rating process by integrating two important physical measures. In this respect, DYFRAT always guarantees an ecological setting for cognitive measurements (e.g., it does not ask respondents to learn what fuzziness is and it works to express judgements/evaluations fuzziness). Moreover, DYFRAT allows to model fuzziness as a natural property which spontaneously arises from some biometric measures (mouse-movements and response times) while respondents use a simple and user-friendly computer device to express their evaluations.
The proposed methodology can be applied in several research fields. For instance, decision making contexts as well as risk assessment situations may require reliable instruments for measuring human uncertainty. Clinical psychologists as well as organizational psychologists may use this instrument for personnel selection problems. However, several other general contexts, such as for instance formative and teaching evaluation, vocational training assessment, decision support systems, quality control and assessment, medical decision making, prevention and treatment assessment, political and social interviews, performance appraisal, military promotion screening, ranking systems, etc., may all require sophisticated tools for adequately measuring the natural uncertainty which arises from these complex situations. We think that DYFRAT can be a reliable, simple and cheap methodology for appropriately addressing such measurement problems.

# Chapter 2

# Representing the dynamics of rating responses: An activation function approach

An extended version of the chapter has been submitted as a research article to *Frontiers in Psychology*.

## 2.1 Introduction

Rating scales are probably the most commonly used measurement tools adopted in education, psychology, social science, and health research because they are flexible scaling procedures for measuring attitudes, opinions, and subjective preferences (Göb, McCollin, and Ramalhoto, 2007; Miller and Salkind, 2002; Aiken, 1996; Pettit, 2002). A rating scale typically consists of a variable to be measured and a set of *anchor points* from which the rater selects the most appropriate description. Among the rating scales, the Likert-type scales are the most widely used scaling methods in education and social science. The main assumption for a Likert-scale is that the strength/intensity of the evaluation is linear, i.e. on a continuum from strongly agree to strongly disagree, with the neutral point being neither agree nor disagree.

Although rating scales are generally as reliable and valid as more complex types of scaling methods (Nunnally, 1978), over the years several criticisms have been arisen against some well-known limitations with this

simple measurement approach. For example, because of the discrete and crisp nature of their format, some individuals tend to avoid extreme categories in the scale (central tendency or restriction range problem) while selecting the final response (Domino and Domino, 2006). Moreover, in some empirical situations (i.e., personality inventories and attitude ) the honesty assumption, tacitly accepted in the administration of self-report rating scale, appears to be simply unrealistic and, therefore the measurements may result in biased observations (e.g., Furnham, 1986). Finally, the standard rating scale paradigm often regards human rating as a discrete-stage based process in which the final response represents its final stage only. Unfortunately, the observed final response simply captures the outcome of the rating process while the real-time cognitive dynamics that occur during this process are usually lost. In particular, the standard observable measures generated during a rating task, the final discrete response and its associated response time, are simply end products of the underlying process of rating, not online measurements of it. In other words, they are indicators of the raters' overall performance, but what really happens during a rating trial is clearly beyond their scope. However, understanding how mental representations unfold in time during the performance of a rating task could be of relevant interest for many researchers working in different empirical domains. Moreover, parsing a rating task into a sequence of subcomponents can help in constructing more sensible indices to detect effects which would otherwise be missed using the standard overall performance measures.

In order to overcome these limitations, here we propose a new family of measures which are designed to track some real-time mental processes unfolding during a rating evaluation. These measures are based on the so-called *mouse tracking methodology* (MTM) and allow to record some relevant cognitive information inferred from the motor control of the computer-mouse during online rating thus capturing temporal and dynamic changing course of data information (Morein-Zamir et al., 2006; Johnson et al.,

2012; Magnuson, 2005; Freeman, Dale, and Farmer, 2011; Jansen, Blackwell, and Marriott, 2003; Hwang et al., 2005; Chen, Anderson, and Sohn, 2001; Mueller and Lockerd, 2001; Freeman and Ambady, 2010; O'Reilly and Plamondon, 2011).

The new measures are assumed to be *observable indicators* of the dynamic process of rating which constitute the antecedents of the final rating outcome and they will allow a) to decompose the observed total rating time into a sequence of temporal subcomponents such as, for example, initiation time, pause time, verification time, and submovement time b) to represent the final response in terms of an activation value which measures the level of intensity/strength for that response. Finally, in our approach both the components are integrated into a common functional model which allows to express the combined temporal and intensity levels of rating.

The remainder of this chapter is organized as follows. In the second section we present the overall idea underlying our approach and provide motivations to use it in human rating problems. In the third section we present our new methodology. In the fourth section we show two empirical applications to real data, whereas in the fifth section we conclude this chapter by providing some final comments.

## 2.2 Rating evaluations as dynamic activation processes

The basic assumption is that identification and exploration of movement phases can provide illuminating features about some cognitive processes involved in the online rating behavior. To introduce the overall scheme of a dynamic rating trial, consider as follows. Let us suppose that a rater is presented with a pseudo-circular scale with $K$ levels (e.g., $K = 5$ for a five anchor-points Likert scale about an item description or question) which appears on a computer screen while the mouse cursor is allocated

to the center of the screen. The rater is asked to provide a response by mouse-clicking the chosen level of the scale (the selected anchor point). Meanwhile, the streaming of the x-y coordinates of the computer mouse (at a given sampling rate) as well as the time sequence of the movements are recorded and stored in the computer memory. The main idea of the dynamic rating framework is to represent each anchor point in the rating scale by a dynamic *activation state*, which indicates (at each recorded time) the level of activation of that anchor point for the current mouse position in the movement path. In general a level of activation at a given instant in time can be understood as a measure of the intensity/strength for the potential final response. The underlying assumption is as follows: the more the mouse pointer approaches the position of a selected anchor point in the pseudo-circular scale, the more the state of the corresponding rating response will be activated. This framework entails a competing activation system where each anchor point competes with the others for the final response. When the mouse pointer is located at the starting position (center of the screen), all the $K$ distinct anchor points will be equally activated at a certain baseline level. However, once the rater starts to use the mouse pointer and moves it in the two-dimensional space of the pseudo-circular scale, the anchor points with a shorter (Euclidean) distance from the current mouse position will start to show larger activation values. By contrast, the anchor points with a larger distance from the current mouse position will decrease their activation states. The proposed framework aims to represent for each potential final response its dynamic activation as a function of the temporal pattern of movements and the distance from the target anchor point position. Figure 2.1 shows a general diagram of mouse-movements with the associated activation functions.

### 2.2.1 Temporal and activation state measures

One important aspect in analyzing movement control is the identification and exploration of several distinct phases of movements (e.g., Pew and Rosenbaum, 1988). Some examples of dynamic activation functions are

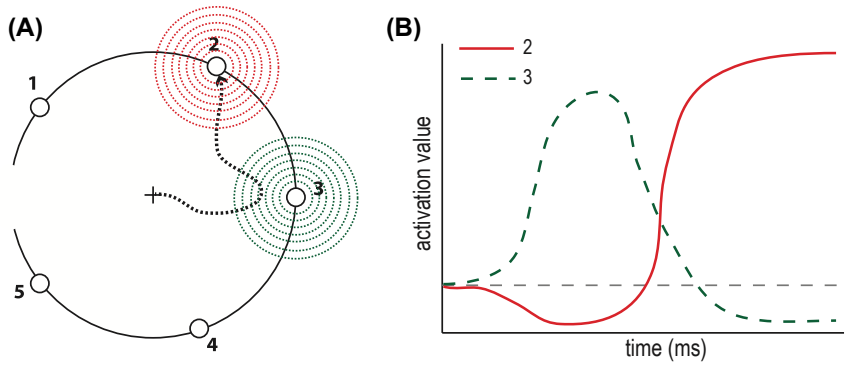



FIGURE 2.1: Hypothetical movement path with activation distributions and corresponding components of positioning time movements. Final response: anchor point 2.

shown in 2.1. In particular, it shows a pattern of movements in which the spatio-temporal uncertainty is related to the rater's choice between two possible alternatives. Note that the associated activation function for the final response has a more irregular shape characterized by two main peaks and one valley indicating the maximal activations for the options 3 and 2 and the change of direction from 3 to 2, respectively.

We can derive a number of quantitative measures about the spatially constrained movements involved in the process of rating of the final selected response. In particular, we can distinguish three different features that play a relevant role in describing the underlying dynamic processes. The first feature is represented by the *static* components of an activation function. These correspond to the portions of the function that are characterized by a flat behavior (see $\beta = 0$ in Figure 2.2-B). The static components are associated with motor pauses in the rating behavior and can reflect the presence of cognitive processes involved in the preparation of the response selection (goal formation) or even in the preparation of a secondary submovement (goal reformulation) for a new selected response option. The second feature refers to the *positive* components of the activation function. Unlike the static portions of the curve, the positive ones are characterized by a positive slope in the activation function (increasing activation function, see $\beta > 0$ in Figure 2.2-B). Positive components denote active movements in the direction of the final selected target and

usually correspond to fast movement executions. Finally, the third feature describes the *negative* components, that is to say, the portions of the function that are characterized by a negative slope (decreasing function, see $\beta < 0$ in Figure 2.2-B). Negative components denote active movements in the opposite direction of the final selected target. These usually reflect temporary deviations from the final response and are also characterized by fast movement executions. In particular, we assume that for executing such submovements further information is taken into account by the rater (e.g., goal reformulation) and that in some circumstances feedforward information may be processed on *the fly* during the same movement production.

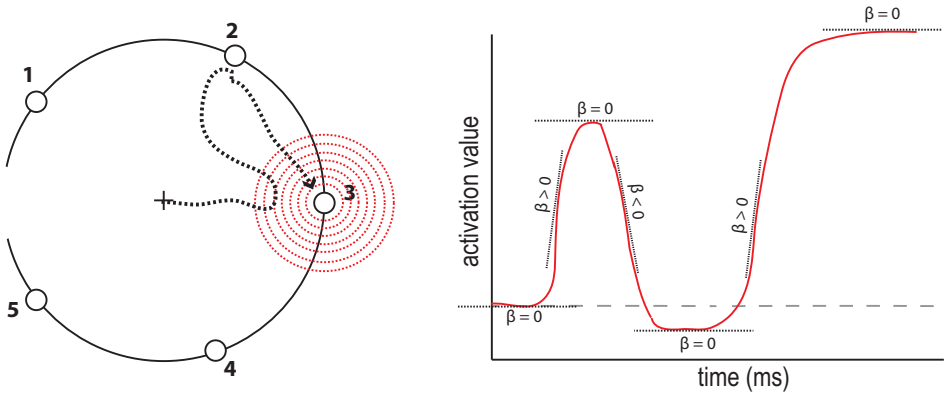



FIGURE 2.2: Hypothetical movement path, activation distributions and corresponding dynamic components.

It is important to note that we can consider each of these functional features as events that are associated with a specific value of the time argument of the function (see Figure 2.3). That is to say, the three features are characterized also by a temporal location.

So, the static components correspond to *pause times* of the positioning mouse movements reflecting temporal phases associated to either goal formation or goal reformulation or response selection (e.g., Kerr, 1978). In particular, the first pause time is associated with the *initiation time*, which represents the time from the stimulus onset and the first recorded movement (overt motion begins). By contrast, the last pause time is connected with the *verification time* which represents the amount of time that a rater

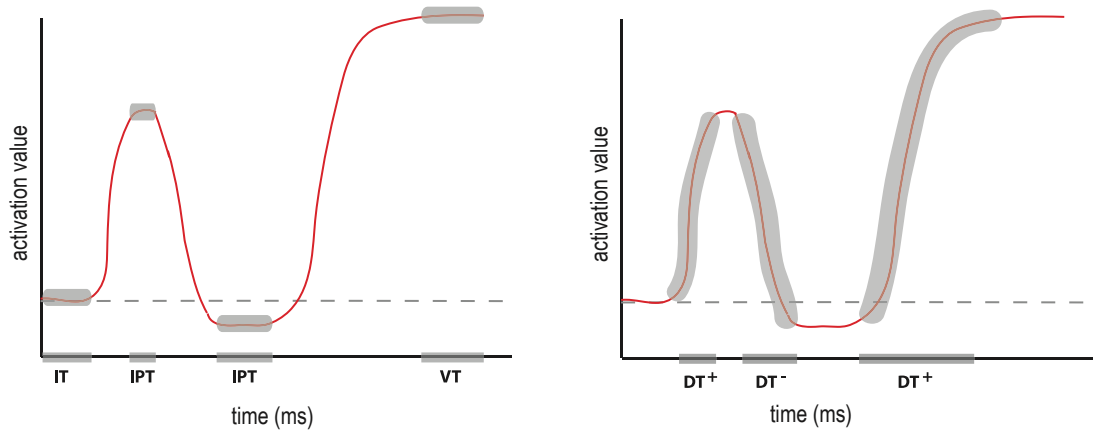



FIGURE 2.3: Quantitative time indices for summarizing the dynamics of the rating process: response time decomposition.

takes to check the location of the cursor and release the mouse button after the last movement has ended. Finally, the positive components (resp. negative components) are associated with *positive submovement times* (resp. *negative submovement times*) in the rating process. These subcomponents allow us to additively decompose the *total time* of the rating process into different parts, each part denoting a different duration for a specific process involved in the dynamic rating behavior. The following decomposition rule for the response time (RT) holds:

$$\mathrm{RT} = \underbrace{\mathrm{IT} + \mathrm{IPT} + \mathrm{VT}}_{\text{pause time}} + \underbrace{\mathrm{DT}^{+} + \mathrm{DT}^{-}}_{\text{dynamic time}} \tag{2.1}$$

where:

- IT is the *initiation time* (from scale onset to first movement)
- IPT is the *intermediate pause time*
- VT is the *verification time* (from last movement to final clicking)
- $\mathrm{DT}^{+}$ is the *positive dynamic time* (time spent toward the target)
- $\mathrm{DT}^{-}$ is the *negative dynamic time* (time spent away from the target)

In general, a high number of pauses in each trial and/or long pause durations will increase the total rating time and can be indicators of some level

of uncertainty in the rating process. Moreover, a high number of submovements (positive or negative) may reflect selection or choice related difficulties for the final rating response option.

In a similar way, we can also decompose the area under the dynamic activation function into subareas according to the temporal phases described earlier (see 2.4).

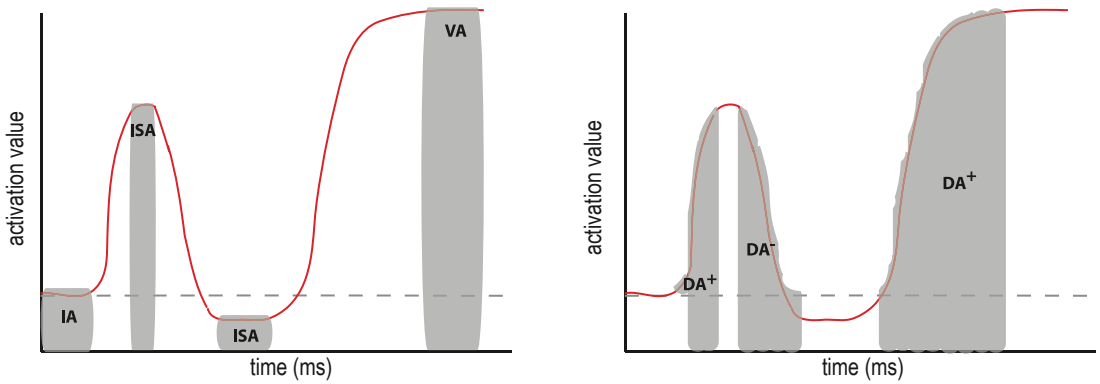



FIGURE 2.4: Quantitative time indices for summarizing the dynamics of the rating process: total activation decomposition.

In general, the area under the activation function indicates the level of intensity or strength for the final selected response. However, like for the times, also for the activation values the different subareas represent separate aspects of the rating process. So, for example, to measure the overall strength for the final response we must consider the sum of the activation integrals associated with the positive submovements and eventually the time spent near the location of the final response (e.g., verification time). By contrast, the subareas associated with the negative submovements should indicate the amount of residual processes linked to temporary deviations from the final response. The total activation (TA) decomposition is as follows:

$$\mathrm{TA} = \underbrace{\mathrm{IA} + \mathrm{ISA} + \mathrm{VA}}_{\text{static activation}} + \underbrace{\mathrm{DA}^{+} + \mathrm{DA}^{-}}_{\text{dynamic activation}} \tag{2.2}$$

where:

- IA is the *initiation activation*
- ISA is the *intermediate static activation*
- VA is the *verification activation*
- $DA^{+}$ is the *positive dynamic activation*
- $DA^{-}$ is the *negative dynamic activation*

Note that quantitative time indices and quantitative activation indices span different information for the rating process. So, for example, we may have two distinct movement pauses (e.g., initiation time and verification time) both characterized by the same interval length but with different overall activation values.

In sum, by using this simple functional framework, we can derive a number of quantitative indices to provide predictions concerning various aspects of the processes involved in the decision mechanism of a rating behavior.

## 2.3 Methodology

Our proposal consists of a *data-capturing procedure* which implements a MTM based computerized interface for collecting the motor and temporal components in the process of rating and a *data-modeling procedure* which provides a functional model for the recorded information. Note that the interface for the data-capturing has been extensively described in the Chapter 1 (sections 1.4.1 and 1.5).

### 2.3.1 Data representation

#### Spatio-temporal data

Let $\mathbf{p} = (\mathbf{x}, \mathbf{y})$ be the *movement path* with length $H + 1$ associated to the streaming of x-y Cartesian coordinates of the computer mouse movements recorded during the rating process (with $H$ denoting the total number

of recorded movements[1]). We assume that in $\mathbf{p}$ the first position $p_0 = (x_0, y_0)$ corresponds to the origin $(0, 0)$ (called *starting position*) of the two-dimensional Cartesian plane $\mathbb{R}^2$, whereas $p_h = (x_h, y_h)$ denotes the $h^{th}$ element (with $h = 1, \ldots, H$) in the sequence of positions recorded in $\mathbf{p}$. In particular, $p_H = (x_H, y_H)$ represents the *final position* in the path and usually corresponds to the position in the two-dimensional plane of the final selected anchor point. Moreover, each position $p_h$ in the path is also associated to a positive integer value $t_h \in \mathbb{N}$ denoting the *time* passed from the onset time of the pseudo-circular scale on the screen and the $h^{th}$ movement recorded in $\mathbf{p}$. By definition we set $t_0 = 0$ (*initial time*). Finally, the total response time (or *final time*) $t^*$ is defined as the difference between the time at the mouse-clicking on the selected anchor point (*final response*) and the onset time of the pseudo-circular scale on the screen. In sum, the spatio-temporal sequence $((p_0, t_0), (p_1, t_1), \ldots, (p_H, t_H))$ constitutes the entire information collected during a single rating trial and the array $(\mathbf{p}, \mathbf{t})$ is the corresponding *spatio-temporal* data structure.

**From spatio-temporal data to functional data**

The main assumption of our approach is to represent each anchor point in the rating scale by an *activation state*, which indicates (at each recorded time $t_h$) the level of activation of that anchor point for the current mouse position $p_h$ in the movement path. To model the activations we used a descriptive perspective based on the simple analogy with bivariate normal densities. In this context, the bivariate normal densities act as sensors to detect mouse movement positions. In particular, let $f_k(\cdot|\mu_k, \mathbf{\Sigma})$ be a bivariate normal density with $\mu_k$ and $\mathbf{\Sigma}$ denoting the location parameter and the scale parameter of the distribution, respectively. The location parameter $\mu_k$ indicates the position in $\mathbb{R}^2$ of the $k^{th}$ anchor point in the pseudo-circular scale, whereas $\mathbf{\Sigma}$ is a diagonal covariance matrix with a single

[1]The value of $H$ can vary trial by trial depending on the dynamic of the recorded rating process.

parameter $\sigma_1 = \sigma_2 = \sigma$ (the parameter sigma is called the anchor point *sensitivity*). Each spatio-temporal observation $(p_h, t_h)$ in the movements path is associated to a positive real value $a_h^k \in \mathbb{R}^+$ denoting the activation of the $k^{th}$ anchor point (with $k = 1, \ldots, K$) for the position $p_h$ recorded at time $t_h$. More precisely, the activation value is given by the following equation:

$$a_h^k = f_k(p_h|\mu_k, \sigma), \qquad h = 0, \ldots, H; \quad k = 1, \ldots, K. \tag{2.3}$$

Finally, the sequence $(\mathbf{t}, \mathbf{a}^k) = ((t_0, a_0^k), (t_1, a_1^k), \ldots, (t_H, a_H^k))$ is the *activation data* structure for the $k^{th}$ anchor point derived from the original spatio-temporal data $(\mathbf{p}, \mathbf{t})$. The basic idea is to think of the observed activation data as a single function instead of a simple sequence of individual observations. Consequently, the activation data can be understood as functional data. More precisely, the term *functional* refers to the intrinsic structure of the data rather than to their explicit form (Ramsay and Silverman, 2005). Note that the activation functions are characterized by the following features: (a) because of the bivariate normal representation, the activation value is nonlinearly related with the distance between the current mouse position and the target anchor point position in the scale (b) the activation functions share all the same sensitivity value $\sigma$ (c) the distance from an anchor point position $\mu_k$ and the starting position $(0, 0)$ is held constant for all the $K$ anchor points in the scale (d) the distance between two consecutive anchor point positions, $\mu_k$ and $\mu_{k+1}$ $(k = 1, \ldots, K-1)$ is also held constant. Note that condition (a) reflects the idea that the activation system is generally less sensitive to movements located far away from the core spots represented by the anchor points' positions. Moreover, conditions (c-d) entail that the positions of the $K$ anchor points lie on a pseudo-circular structure centered on the origin $(0, 0)$.

### 2.3.2 Data modeling and parsing

In our approach we assume that the curve being estimated is continuous and smooth. There are several methods that can be considered for approximating discrete data by a function (for a review see Ramsay, 2006). In this contribution we adopted the *roughness penalty* or *regularization* approach which is a powerful and flexible option for modeling functional data based on cubic B-spline basis functions (for more details the reader may refer to Ramsay and Silverman, 2002). Figure 2.5 shows the approximation of the observed movement data represented in the panel A with a smooth activation function (panel B).

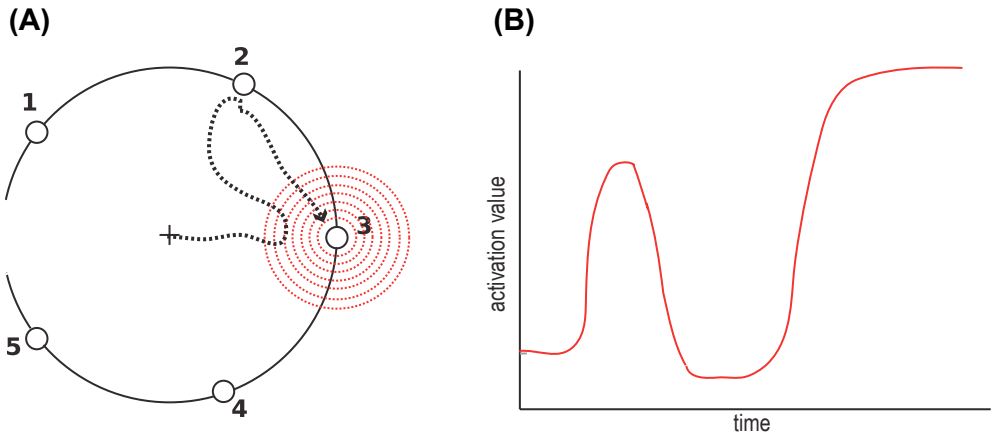



FIGURE 2.5: Hypothetical movement path (A) with the associated smooth activation function (B).

One of the main advantages of modeling discrete data by a functional data analysis approach is that we can easily derive smooth approximations of the original activation data which may instead be characterized by not so regular representations due to the limitations of the empirical sampling rate.

Once the activation function and its first derivative have been modelled using the functional data approach, we need a procedure for parsing the activation function into its basic components. In particular, we extract the dynamic components and static components in order to derive the quantitative indices for the rating process described earlier. We used a simple

parsing algorithm adapted from previous analytical techniques to detect corrective submovements. Our approach entails a systematic evaluation of the first derivative (velocity) of the activation function. In this evaluation, submovements as well as pauses are defined on the basis of certain criterion events (e.g., crossing of well-defined velocity thresholds). These criteria are chosen to take account of differences between the dynamics of voluntary movements (Meyer et al., 1988), physiological tremors and passive residual activity due to springlike characteristics of muscles. In particular, a static component of the activation function representing a movement pause was defined to be a continuous portion of the function such that the absolute value of its first derivative did not exceeded a given small positive threshold $\delta$ and the duration of the corresponding pause exceeded 20 ms.

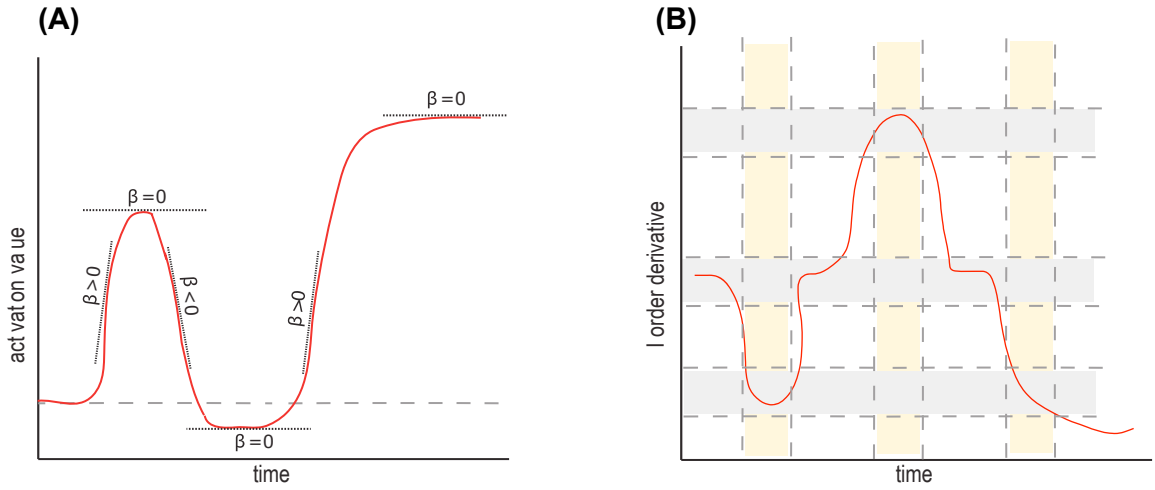



FIGURE 2.6: Hypothetical smooth activation function (A) with the corresponding first-order derivative function (B). Note that dotted lines refers to a parsing grid.

The conjunction of these two conditions defined the criterion to extract aimed pauses from the dynamic rating process to detect temporal phases associated to either goal formation or goal reformulation or response selection. Accordingly, a dynamic component (either positive or negative) of the activation function representing an aimed submovement was defined to be that portion of the function in between two consecutive pauses.

The latter means that a positive submovement (resp. negative submovement) must show activation values with positive first derivatives exceeding $\delta$ (resp. with negative first derivatives not exceeding $-\delta$) and remained above (resp. below) that level continuously for at least 20 ms. These criteria helped ensure that the putative submovements stemmed from voluntary activity rather than physiological tremor or passive damped oscillations between two submovements. Figure 2.6 shows the adopted parsing rationale.

## 2.4 Illustrative examples

By way of illustration, in this section we describe two simple empirical applications using our methodology. The first example considers a well known cognitive problem in decision making called the *trolley scenario*. The second application presents data about the simulation of faking responses in binary categories. We stress that the examples reported here have only illustrative purposes and are presented in order to highlight some important aspects of modeling the temporal and activation state measures described earlier.

### 2.4.1 Rating responses and moral dilemma

*General context and motivation*. Moral judgements and dilemmas are important puzzles studied in cognitive decision making Greene and Haidt, 2002; Haidt, 2001. Because of their complex nature and representation, moral dilemmas are usually characterized by some level of uncertainty in individuals responses. One important example of moral dilemma is the so called *trolley scenario* Greene and Haidt, 2002; Haidt, 2001; McGuire et al., 2009; Nichols and Mallon, 2006. In general, individuals react to this dilemma by using their personal moral beliefs. More precisely, the trolley scenario engages some levels of emotional processing which, in turn, might affect people's judgments to such an extent that some individuals may definitively hesitate in evaluating the scenario. Some authors (e.g.,

Haidt, 2001; Bruner and Bruner, 2009) claim that moral evaluations can be distinguished in two distinct categories: moral intuition and moral reasoning. Generally, moral intuition occurs quickly, effortlessly, and automatically, such that the final judgment (but not necessarily the underlying process) is accessible to consciousness. By contrast, moral reasoning occurs more slowly, requires some effort, and involves at least some steps that are accessible to consciousness. In this first example, we assessed the extend to which each temporal measure is affected by the type of moral evaluation as codified by the rater's response[2]. Moreover, we also studied the effect of gender on these temporal measures. We expect that the individuals who show a very strong disagreement with the action described in the trolley scenario (moral intuition raters) will be characterized by faster responses (according to one of the temporal indices) with low levels of uncertainty. By contrast, those who are characterized by a more moral thinking attitude (moral reasoning raters) will show less extreme responses with larger values of uncertainty as measured by slower response and more complex movement patterns. Because the trolley scenario may activate not necessarily conscious underlying mechanisms in the rater, we believe that our methodology can represent an ideal tool for studying the subcomponents of the online rating process.

*Participants*. One hundred and tree volunteers from the university of Trento (males = 47, age 18-23 : $70.87\%$, age 24-27 : $19.42\%$, age 28-36 : $2.91\%$, age $\geq 37$ : $3.88\%$) were tested after providing written informed consent. All participants had normal or corrected-to-normal vision. The study was approved by the local ethics committee.

*Moral dilemma*. In this application, we used a moral dilemma based on the well-known *trolley scenario* Greene and Haidt, 2002; Haidt, 2001; McGuire

[2]For the sake of clarity, in this first application we limit our analysis to the temporal indices only.

et al., 2009; Nichols and Mallon, 2006 as described in Chapter 1, section 1.6.1.

*Procedure*. See Chapter 1, section 1.6.1.

*Movement parsing*. At the end of the data collection, the records of the participants' positioning movements and movement times were first modelled (separately for each participant) using the functional approach described earlier and subsequently parsed to isolate component submovements and pauses (see Section 2.3.2). In particular, in modeling the functional data we adopted a smoothing parameter $\lambda = 1.5$ for all the 103 participants. After submovements were isolated, the total rating time on each participant was separated into the initiation time, verification time, total positive dynamic time, total negative dynamic time, and total pause time.

*Data-analysis*. Because the temporal measures considered in this application showed a nonzero level of skewness and kurtosis, we preferred to analyze our data by using robust data analysis procedures. In particular, we performed robust two-way between-subjects analyses of variance on each of the dependent variables mentioned previously, with *moral type* (*M*: moral intuition rater vs moral reasoning rater) and *gender* (*G*: male vs female) as independent factors at two levels each. Note that in order to define the moral type factor we dichotomized the observed discrete rating responses into two distinct levels: $1$ vs $> 1$. The idea behind this dichotomization was that moral intuition raters would be characterized by the most extreme response: *strongly disagree* whereas the moral reasoning raters would prefer less extreme responses (e.g., $> 1$). The robust analyses of variances were performed using the `robust` package available in the `R` environment.

*Results*. Figure 2.7 shows the results of this first application. About the

temporal measures, mean $RT$ increased in the male group ($p < 0.05$). However, the main effect of factor M and the interaction M by G were not significant. In order to interpret these results the durations of the distinct movement phases must be examined. Mean $PT$ for moral intuition raters was significantly shorter than the corresponding mean $PT$ for the moral reasoning raters ($p < 0.05$). A similar effect was observed also for the gender factor with the female group showing shorter $PT$ than that observed for the male group ($p < 0.01$). We did not observe any significant interaction between the two factors. However, looking at the decomposition of the total pause time reveals some interesting differences between the two factors M and G. In particular, mean $IT$ for moral intuition raters was significantly shorter than the corresponding mean $IT$ for the moral reasoning raters ($p < 0.01$). By contrast, the mean $VT$ for the female group resulted shorted than that of the male group ($p < 0.05$), whereas all the other effects were not significant for this variable. Finally, no other statistically significant differences were observed for the remaining temporal measures ($DT^+$ and $DT^-$).

*Discussion*. The results of this first application suggest two complementary conclusions. First, to provide a substantive interpretation of dynamics of a rating task, it is not sufficient simply to examine the total rating time in order to characterize how positioning movements and absence of movements (pauses) are produced. Rather, an analysis of the component phases that contribute to these times, including the durations of movement initiation, execution, and verification, must be made. Second, initiation times and verification times are affected somewhat differently by the same independent variables. In particular, mean $IT$ for moral intuition raters was shorter than the mean $IT$ for the moral reasoning raters. Because $IT$ represents the time from the stimulus onset and the first recorded movement (overt motion begins), we may consider this temporal measure as directly correlated with the time spent by an individual in choosing among the $K$ distinct options. In general, a long $IT$ duration can reflect some level of

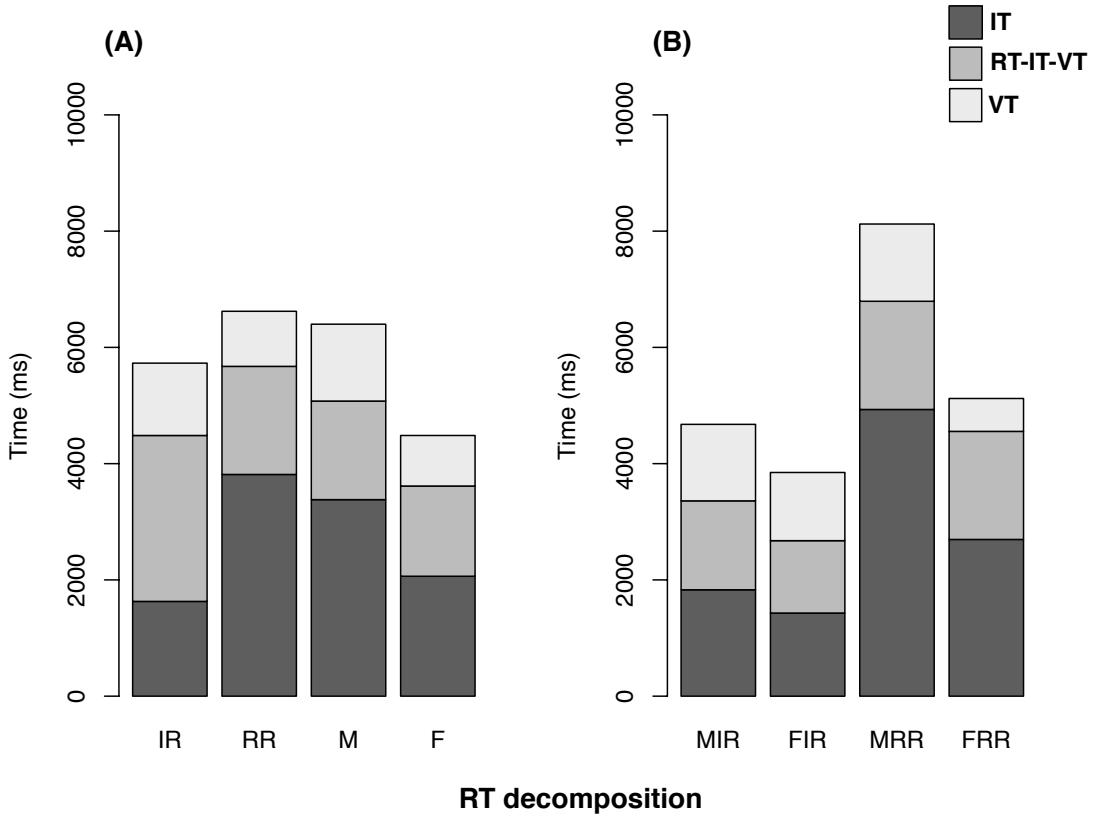



FIGURE 2.7: Application 1: Robust Anova for the RT decomposition. Note that in the panel A the acronyms IR means *initiation rate*, RR *rationality rate*, M stands for *male group*, and F for *female group*. Instead in the panel B the acronyms refers to the interactions among the factors (e.g., MIR means interaction between M and IR).

uncertainty in selecting the preferred option among the alternatives. In this respect, moral reasoning raters seem to show a relative larger level of uncertainty as compared to moral intuition raters. This result is in line with the theoretical expectation that moral intuition raters are characterized by more extreme, quick, effortless, and automatic final responses. By contrast, because $VT$ measures the amount of time that a rater takes to check the location of the cursor and release the mouse button after the last movement has ended, we may understand this temporal measure as correlated with the time spent by an individual to provide a final check for the selected response. In general, a longer $VT$ duration can indicate some level of uncertainty in confirming the final response. Interestingly, in this respect the female group seemed to show less uncertainty in confirming the final response as compared to the male one.

### 2.4.2 The dynamic structure of faking

*General context and motivation*. In real life contexts, some individuals tend to distort their behaviors or actions in order to reach specific goals. Nowadays there is a broad consensus that faking/deception is an intentional response distortion aimed at achieving a personal gain (Ziegler, MacCann, and Roberts, 2011; Vrij and Verschuere, 2013). For example, in personnel selection some job applicants may misrepresent themselves on a personality test hoping to increase the likelihood of being offered a job (Anderson, Warner, and Spencer, 1984). Similarly, in the administration of diagnostic tests individuals often attempt to malinger post traumatic stress disorder in order to secure financial gain and/or treatment or to avoid being charged with a crime (Hall, Thompson, and Poirier, 2007). There are several studies that show how lying is cognitively more demanding than truth telling. According to this approach the additional cognitive load is caused by fakers having to being involved in extra tasks such as, for example, inferring what others are thinking or monitoring their behaviour so that they may still look natural to the other people's eyes. Indeed, research shows that compared with truth tellers, liars are slower and make more errors when responding to questions as it seems that truth is automatically activated in the cognitive system of a person (Walczyk et al., 2005). In sum, faking behaviors seem to require inhibition of prepotent truth responses.

In this second application, we evaluated the performances of the temporal and activation state measures in an empirical test based on simulated faking. We used a questionnaire about episodic memory administered in a laboratory-type situation in which two different experimental conditions (honest motivating condition vs faking motivating condition) were manipulated for the responders. Specifically, we analyzed the mouse tracker patterns that occurred while participants had to perform a binary (yes-no) response task to items of a questionnaire about episodic memory under

honest or faking conditions. By measuring the moment-by-moment temporal and spatial dependencies that reside in mouse-movement patterns of hand motion we are potentially able to provide a dynamical account of fakers' online classification in a pseudo-naturalistic scenario. Moreover, because in this context simple response time may reflect a combination of different subprocesses that potentially differ in sensitivity to the effect of faking, we expect that the decomposition of the temporal and activation state measures can enhance the discrimination between simulated true and fake responses. Finally, in this application we show how the proposed approach can be used to evaluate the dynamics of responses based on binary ratings.

*Participants*. Fifty two participants from the university of Padua (males = 34, age 18-23 : 4, age 24-27 : 17, age 28-36 : 13) were tested after providing written informed consent. All participants had normal or corrected-to-normal vision. The study was approved by the local ethics committee.

*Materials, apparatus and procedure*. The questionnaire was composed by 100 true/false statements referring to episodic memory items. The 100 statements were divided into four experimental conditions: a) true-yes responses b) fake-yes responses c) true-no responses and d) fake-no responses. Each of the four conditions comprised 25 distinct sentences. Some examples of item statements were: *the experimenter welcomed me* (true-yes response); *I am at a rock concert* (fake-yes response); *I am in Paris* (true-no response); *I am using the mouse* (fake-no response). The questionnaire included a male and a female version for the 100 statements.
We used the same apparatus described in the first application. However, the experimental procedure was adapted for fitting specific requirements of this second study. In particular, unlike the fist application, in this new study the participants responded to all the 100 items of the questionnaire. The same randomized item order was presented to all participants and the 100 experimental items were first preceded by five initial control items. To

modulate the experimental factor *instruction type* (honest vs. fake), the colour of the pseudo-circular scale was appositely manipulated to instruct the participants on whether to respond honestly (grey scale) or lying (red scale). The available anchor points in the pseudo-circular scale were S (for Yes, positive response), ? and N (for No, negative response) (see Figure 2.8). Participants were told to answer spontaneously, with no time pressure, and to ignore the question mark (?) option, which acted only as the neutral reference in the scale configuration.

Like for the previous example, also for the analysis of the faking questionnaire the functional approach was used to model the participants' positioning movements and movement times (separately for each participant) with smoothing parameter $\lambda = 1.5$. After the submovements were isolated using the parsing algorithm, all the main temporal and activation state measures were computed for each participant and each item in the questionnaire.

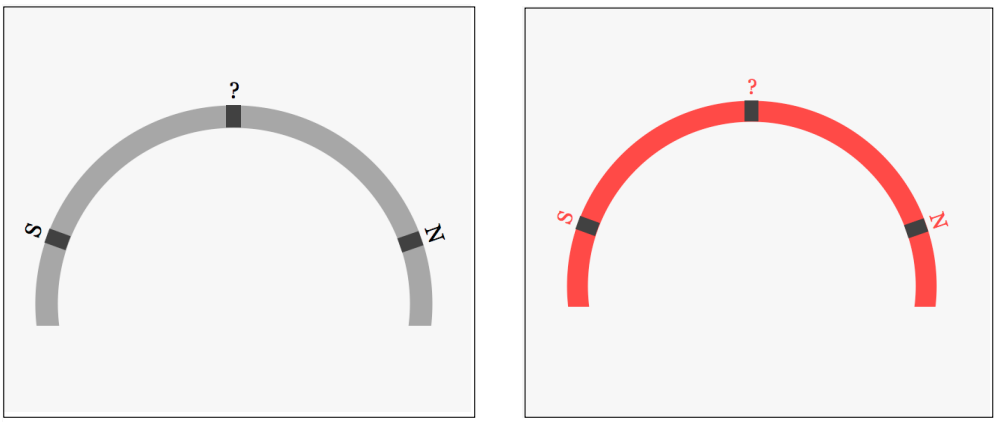



FIGURE 2.8: Application 2: Example of scale used with experimental factor "colour".

*Data-analysis*. The experimental procedure yielded data that could be organized according to a two-way within-subjects factorial design: *response type* (Y vs. N) $\times$ *instruction type* (honest vs. faking). For each participant and each dependent measure (temporal and activation state measures) we

derived four summary statistics based on the median values of the original measures collected at the item level. In particular, the four summary statistics were associated at the four distinct combinations of the two-way factorial design. For the sake of simplicity, in this second study we will analyze the data for the two levels of the response type factor (Y vs. N responses) separately.

*Results*. About the temporal measures, we analyzed the following indices: total response time, overall dynamic time, overall pause time, initiation time, and verification time. For the summary statistics computed on the positive items (Yes-type responses), all the temporal indices showed more or less the same homogeneous pattern (see Figure 2.9). As expected, fake responses were in general characterized by longer response times ($p < 0.001$). In particular, the fakers showed longer dynamic times ($p < 0.05$) and longer pause times ($p < 0.05$) compared to honest respondents. However, we observed subtle differences between the subcomponents of pause time. More specifically, for the fake responses the initiation time was significantly longer than the corresponding time for honest responses, that is to say, fakers spent more time before starting the first movement. By contrast, the two typologies of responses did not show any significant difference for the verification time as they required approximately the same amount of time to check the final response and release the mouse button after that the last movement has ended. However, for the summary statistics computed on the negative items (No-type responses) the behavior of the temporal measures was clearly different (see Figure 2.10). In particular, for no-type responses the overall response time was not significantly different between fakers and honest respondents. Apparently, the bigger cognitive loads associated with faking behaviors disappeared when negative items were considered. Similarly, also the overall dynamic time was substantially similar between the two categories of responses. By contrast, the pause sub-components showed significant differences between fake responses and honest responses. However, this time the observed patterns

were reversed: honest responses had longer initiation times ($p < 0.05$), whereas fake responses were characterized by longer verification times ($p < 0.05$).

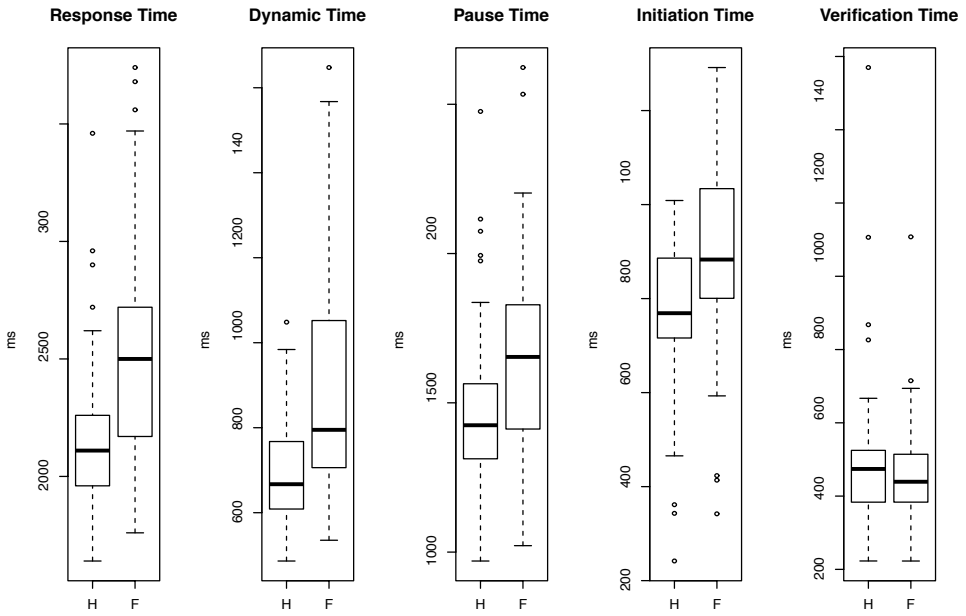



FIGURE 2.9: Application 2: Box-plots for the honest and fake condition in the YES-type responses (RT decomposition).

About the activation state measures, we analyzed the following indices: total dynamic activation, positive dynamic activation, negative dynamic activation, static activation, initiation activation, and verification activation. For the positive items (Yes-type responses), the observed results were very similar to those reported for the temporal measures (see Figure 2.11). All the activation measures but verification activation were significantly different between honest responses and fake responses (all $p$s $< 0.05$). Likewise for the temporal indices, also the activation indices showed an inverse pattern when negative responses were considered (see Figure 2.12). In particular, in the no-type response condition only initiation activation and verification activation resulted significantly different in the two instruction modalities with initiation activation being larger for honest responses ($p = 0.056$) and verification activation being smaller for

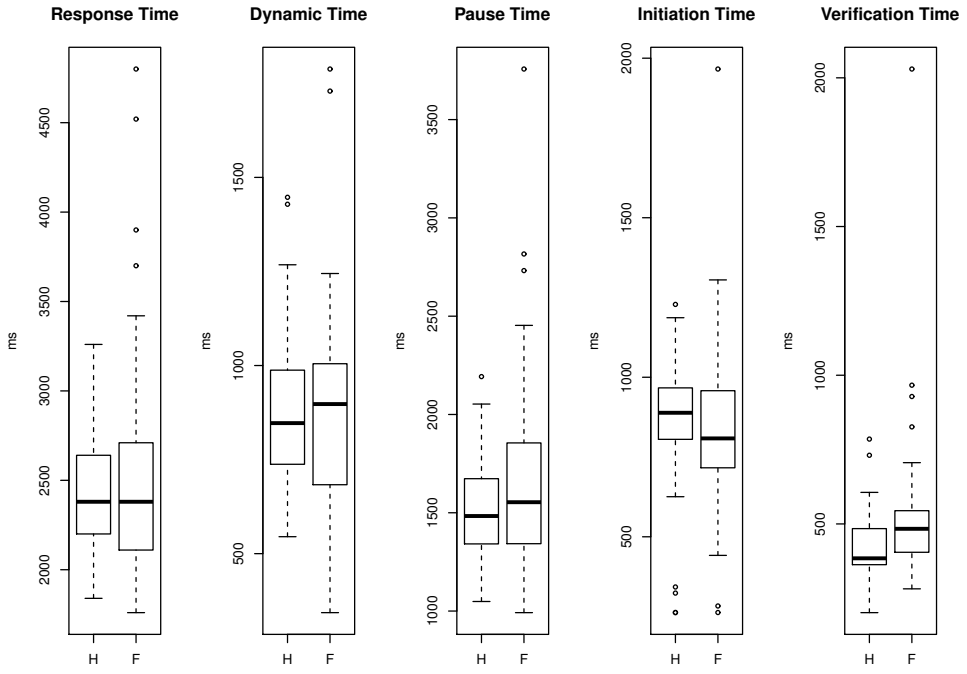



FIGURE 2.10: Application 2: Box-plots for the honest and fake condition in the NO-type responses (RT decomposition).

fake responses ($p < 0.05$). Overall, the activation indices mirrored the results observed for the temporal indices.

*Discussion*. Generally, faking requires more cognitive efforts relative to truth-telling, including truth suppression and consistent fabrications. This means liars constantly deal with a considerable amount of strategies they have to manage in order to get away with their lies. The results of this second application suggests that additional cognitive loads play a relevant role in characterizing the faking pattern for positive responses only. In general, a true yes response requires less time and less effort to be produced compared to a fake yes response. Things are different for negative-type responses, however. Fake no responses and true no responses are substantially indistinguishable according to standard performance measures such as, for example, response time. By the way more subtle measures as verification time (resp. verification activation) or initiation time (resp. initiation activation) may help in discriminating between honest

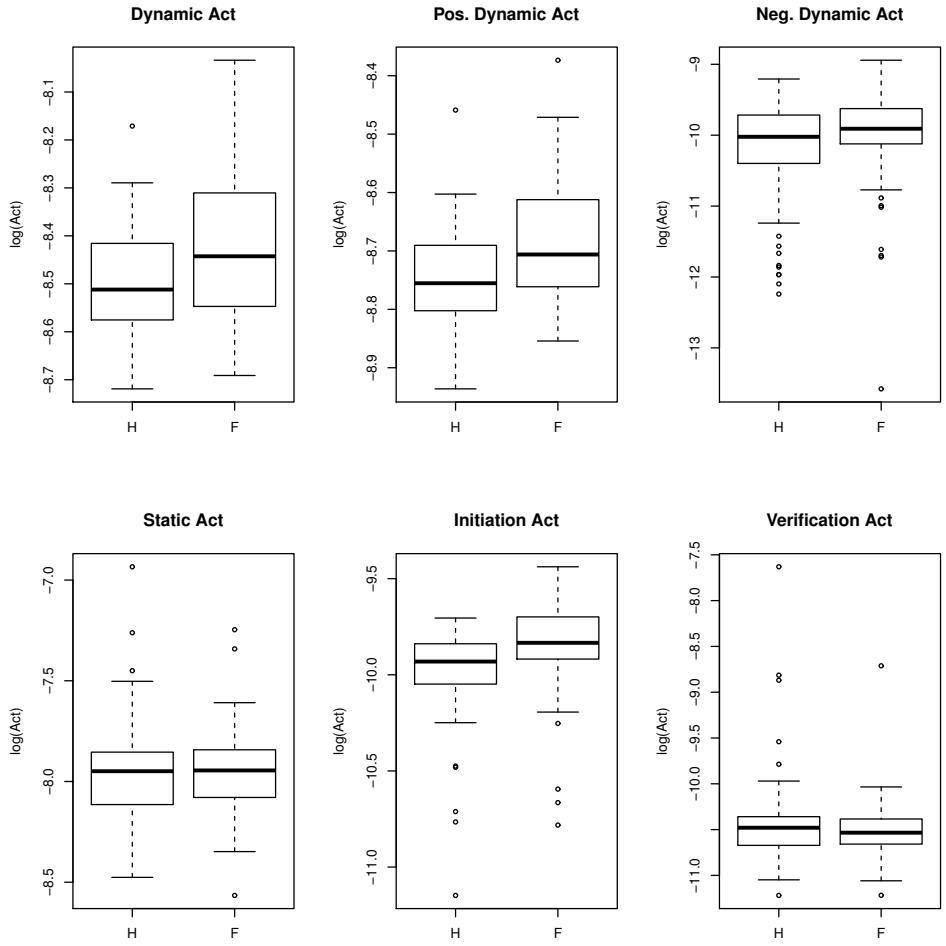



FIGURE 2.11: Application 2: Box-plots for the honest and fake condition in the YES-type responses (TA decomposition).

and fake responses. We speculate that for negative responses the additional cognitive load observed for fake yes responses is somehow absorbed or masked by the additional effort which is normally required to elaborate and produce negative information. In general, negative information seems to require greater information processing resources and activity than does positive information (e.g., Taylor, 1991) and this additional processing leads to differences between positive and negative information in attention, learning, memory, and apparently also in faking.

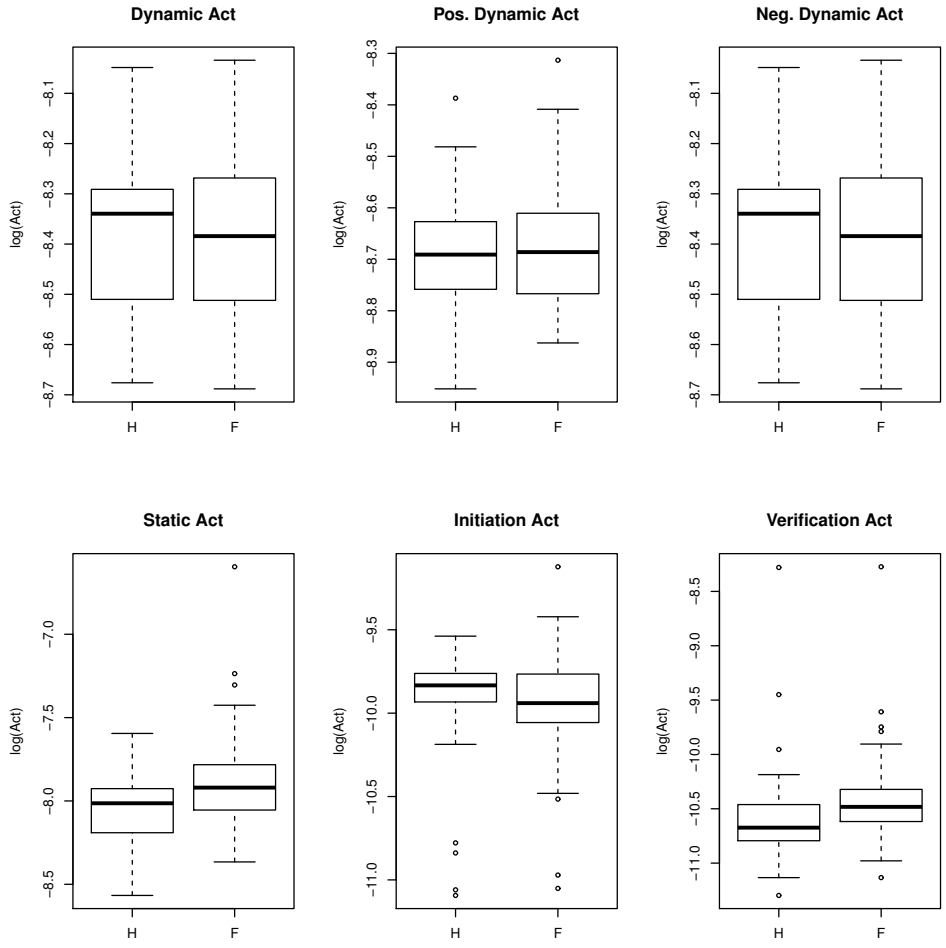



FIGURE 2.12: Application 2: Box-plots for the honest and fake condition in the NO-type responses (TA decomposition).

## 2.5 Conclusions

In the present contribution an innovative approach for measuring the process underlying rating evaluations was presented. The simple final value recorded in a rating task does not provide any insight into how a person makes a decision: did the rater react immediately and quickly or did he/she show some doubts and maybe change his/her mind a couple of times? Here we tried to overcome this fundamental limitation by creating a tool able to capture some hidden components of the response process. By recording of mouse trajectories and inherent response times while subjects express their ratings on selected anchor points of a pseudo-circular scale; we were able to derive biometric physical measures thought to act

as observable indicators of the underlying cognitive processes (Magnuson, 2005). These measures can be used to construct more sensible indices to detect effects which would otherwise be missed using the standard overall performance measures. In two distinct empirical applications we showed the validity of our new framework in cognitive settings.

# Chapter 3

# A simple information theoretic approach for modeling spatial data from mouse tracker methodology

The content of the chapter partially reproduces a research paper which will be considered for publication.

## 3.1 Introduction

Human movement is an integral part of perceptual-motor system. Most people constantly use their arms, hands, and fingers in simple and complex activities, for example, in reaching objects, pointing at visual targets, typing on a computer keyboard or moving a computer mouse. In all these activities, movements can be very complex and their study, in terms of kinematics and/or simple trajectories, can provide a simple and valid basis to uncover cognitive and behavioral processes underlying the observable structure of the data. To this aim, many research groups have focused their attention on the analysis of hand movements and their trajectories, which are very often collected by means of a computer mouse. It has been shown that such device is a simple, reliable, and accurate real-time data acquisition device that can be reasonably adopted for experimental purposes (O'Reilly and Plamondon, 2011; Morein-Zamir et al., 2006). As a consequence, mouse tracking based experiments are becoming very popular in cognitive studies over the recent years (Friedman, Brown, and Finkbeiner, 2013; Hehman, Stolier, and Freeman, 2014). In general, the

main idea of the mouse tracking methodology is to consider the collected movement trajectories as a continuous source of real-time information on the internal cognitive processes activated during a particular experimental task. For instance, in a dichotomous categorization task where two alternative choices are presented on a computer display (e.g., target vs. distractor), the features of the computer mouse trajectories can reveal, according to the experimental manipulation, the eventual competitive attraction that one of the two choices (e.g., distractor) acts on the correct category (target) (Spivey and Dale, 2006). This simple idea has been recently applied in several research, including social categorization (Dale, Kehoe, and Spivey, 2007), moral decision making (Koop, 2013), language comprehension (Morett and Macwhinney, 2013), and numbers representation (Faulkenberry, 2014).

By and large, different strategies can be adopted by researchers in dealing with mouse movement trajectories. Without claiming to be complete we can enumerate four strategies, as follows. The first strategy adopts a *statistical-probabilistic framework* (e.g., see Friedman, Brown, and Finkbeiner, 2013; Selen, Shadlen, and Wolpert, 2012) where, in particular, the movement paths are evaluated by fitting dynamic models on the collected data (e.g., Wiener's Diffusion models, dynamic Markov Chains) and by evaluating their results from a predictive viewpoint. Although this strategy can be very powerful in discovering important features present in the empirical path of movements (e.g., movement initiation process, decision phases, changes of direction), it requires to make several assumptions on the nature of the data, in particular on how the movements are generated. In addition, often such methods can also show some pragmatic difficulties in estimating the models' parameters with the consequence that richness and heterogeneity of empirical movement paths cannot be accurately represented (e.g., see Voss and Voss, 2008; White et al., 2010). The second strategy adopts a *mathematical framework* through which the motor movement is represented in terms of physical models (like the Rayleigh-Duffing model) and kinematic principles (e.g., minimum jerk, minimum energy)

(Plamondon, 1995; Flash and Henis, 1991; Engelbrecht, 2001). Unlike data-driven approaches, here the harm movement is described using a top-down perspective where its main characteristics, such as speed, acceleration, stability, and friction, are formally deducted by the general model's equations. Generally, this makes the approach more precise and accurate in modeling and predicting the motor components of the movement. However, despite that fact, in some circumstances these approaches might be not flexible enough to model structured and noisy movement paths that are usually encountered in real situations (Shadmehr, 2005). The third strategy uses a *computational framework* to decompose the recorded movement path in different sub-components (Walker, Meyer, and Smelcer, 1993; Meyer et al., 1988). The main idea here is that parsing a goal-directed movement into a sequence of informative subcomponents (e.g., initiation time, pause time, verification time, number of sub-movements, task completion time) can help in constructing more sensible indices to detect effects which would otherwise be missed using standard performance measures, such as simple reaction-times (RT). Overall, this strategy carries out considerable information although it might fail in providing global movement information, especially in contexts where, for example, researchers are interested in evaluating the dynamic attraction toward competitive cues (Thompson et al., 2007; Smits-Engelsman, Van Galen, and Duysens, 2002; Hwang et al., 2005). Finally, the last strategy naively makes use of a *descriptive geometric framework* that simply considers the Cartesian coordinates of the movement trajectories as input of further analyses. As a consequence of this choice, simple statistics such as, local means, distances, area under the curves, point-by-point t-tests, are used as typical and basic measures to globally evaluate the recorded movement path (e.g., see Hehman, Stolier, and Freeman, 2014; Freeman et al., 2008; Freeman and Ambady, 2010). Although such approach is simple and cheap to run, very often it can incur considerable shortcomings with a significant loss of information. For instance, its derived measures might not be sufficiently flexible and accurate to capture more complex structures (e.g., quasi-sinusoidal,

multi-peaks, irregular paths) as well as the presence of motor pauses in the movement paths (Wel et al., 2009; Fischer and Hartmann, 2014).

In this paper, we exclusively focus on the descriptive geometric approach and its formal characteristics. The reason of this is that, despite its limits, this strategy of data analysis results to be simple, cheap, and easy to carry out on a large scale of research contexts. Although very often this approach can yield improper results, a considerable number of papers making use of its rationale have appeared over the recent years. Bearing this in mind, in this paper we propose a new mathematical approach based on an *information-theoretic paradigm* which is able to overcome some of the main limitations of the aforementioned approach. In particular, in our approach the empirical movement path is firstly statistically modeled by using an adaptive cumulative function that is able to capture the direction, amplitude, and eventual frictions that are naturally present in the movement. Next, a set of entropy-based measures, which include a movement-pause entropy decomposition, provides an analytic way to quantify the most relevant spatial information that are present in the empirical data trajectories. These new measures are provided to be more sensitive, robust, and stable than the descriptive geometric measures. These results are especially supported by means of specific simulations studies. It should be noted how, unlike the descriptive geometric approach, our approach would also provide an appropriate framework to compare and evaluate different movement patterns by means of suitable distributional distances.

The reminder of this chapter is organized as follows. In the second section we describe a comprehensive survey on the rationale underlying the descriptive geometric approach and the derived measures and indices. In the third section we present our proposal together with its main characteristics. The fourth section illustrates some simulation studies carried out on our methodology whereas the fifth section concludes this chapter by providing some final comments and suggestions for future extensions of

our approach.

## 3.2 Descriptive geometric approach: a brief overview

### 3.2.1 General context and motivation

The descriptive geometric approach (DGA) has been widely used to discover the cognitive features associated to the observable motor behaviors of the hand. In particular, some of the earliest research used DGA as strategy to analyze mouse tracking data from choice, categorization, and decision-making experimental tasks (Song and Nakayama, 2009). To give an example, consider a two-choices categorization task where a subject is asked to select, by means of a computer mouse, the appropriate category of a stimulus against a competitive alternative (Figure 3.1). Figure 3.1-A shows the conceptual categorization schema through which a given stimulus presented on the display (e.g., picture of a *dolphin*) has to be assigned to the right target category C1 (e.g., *mammal*) against the competitive alternative category C2 (e.g., *fish*). The underlying hypothesis here is that the movement of the hand, which is registered with a computer mouse (in green color in the Figure 3.1-A) together with the associated continuous streams of motor information, might reveal important aspects of the ongoing cognitive competitive dynamics involved during the task (e.g., the more the similarity between the categories *mammal* and *fish*, the more the difficult in the categorization task, the more the curvature toward the incorrect category *fish*).
Figure 3.1-B/C shows the geometric information in the two-dimensional Cartesian plane which is associated to the motor movement path and, in particular, the panel (C) shows two basic and important descriptive measures used in this approach, namely the maximum deviation (MD) and the area under the curve (AUC).

Generally, DGA basically implements a two-steps strategy of data analysis. In particular, to obtain a set of consistent $x$-$y$ trajectories, the raw

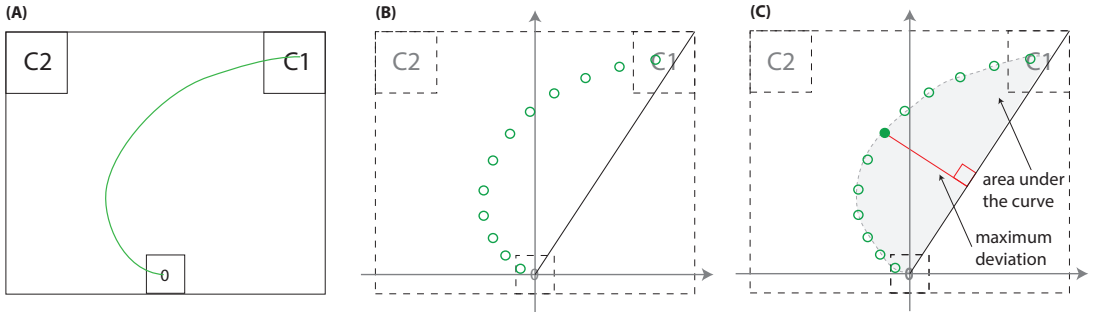



FIGURE 3.1: Conceptual diagram of DGA in a two-choices categorization task. (A) Categorization schema with target C1, competitive category C2, and an hypothetical mouse path in green color. (B) Cartesian mouse trajectories in 2-dimensional plane (the bisector of the I quadrant indicates the ideal movement). (C) Example of two descriptive measures of the movement path: gray color indicates the area under the curve (AUC) whereas the red straight line indicates the maximum deviation/distance (MD).

computer-mouse trajectories that are recorded during a cognitive task are firstly roughly pre-processed by space rescaling, time-normalization, and averaging. Lastly, the ensuing movement paths are carved up by using various measures (e.g., MD, AUC, *x*-*y* flips, sample entropy, bimodal coefficient) that are supposed to be good indicators of the ongoing dynamics involved in the cognitive tasks. These final measures are finally used in running post-hoc specific analyses, such as the so-called spatial attraction analysis.

### 3.2.2 Procedures and data analysis

For the sake of description, all the data analysis procedures involved by DGA can be grouped into four main categories, namely spatial attraction analysis, velocity-acceleration analysis, distributional and complexity analyses (Hehman, Stolier, and Freeman, 2014; Freeman and Ambady, 2010). In particular, *spatial attraction analysis* is mainly performed by means of AUC and MD measures. These indices are computed after space rescaling, time normalizing, and averaging the raw coordinates by means of participant's mean profiles and grand mean profiles. Technically, participant's mean profiles are obtained by averaging the raw *x*-*y* trajectories over the trials for each individual present in the dataset whereas the grand

mean profile is carried out by averaging together all the data from participants and trials. Finally, MD can be computed over such profiles as the largest perpendicular distance between the averaged trajectory and the ideal movement path linking the starting point position to the correct cue (see Figure 3.1B-C). Similarly, AUC is defined as the area between the mouse trajectory and the ideal straight movement and can be computed through any numerical integration method (e.g., trapezoidal method). Note that, both MD and AUC measures would assess to what extent the magnitude of activation of the competitive cue changes under varying movement trajectories. Particularly, "AUC is a better index of the overall attraction toward the unselected alternative whereas MC is a better index of maximum attraction" (Freeman and Ambady, 2010, p.230). The *velocity-acceleration analysis* is conducted after rescaling raw trajectories with an user-defined number of pre-fixed time bins. As for the previous analysis, also velocity and acceleration profiles can be obtained for each participant's trajectories and/or for each group of participants/experimental condition. In particular, velocity is roughly computed by taking the distance between couples of $x$-$y$ coordinates at different times whereas acceleration is obtained by measuring the relative changes in velocity across time points. Unlike the previous analysis that are conducted on the raw trajectories collected over the experimental tasks, the *distributional analysis* applies on the final spatial measures (e.g., AUC). Technically, all the computed AUCs over the participants and conditions are firstly $z$-normalized and lastly represented as a discrete (binned) distribution. The rationale is to analyse whether the distribution is unimodal or bimodal (bimodality is usually assessed by means of Sarle's $b$-test). Unlike unimodal distributions, a bimodal distribution would always indicate the presence of two sub-populations showing two different type of movement responses and, therefore, can be used in assessing whether attractions toward competitive cues occurred. Finally, the so-called *spatial*

*disorder analysis* can be employed to assess the "complexity" of mouse trajectories that would reveal how much irregular, unpredictable, and disorganized, the mouse movements appear over the experimental conditions. Particularly, such an analysis is performed by computing (i) the number of reversal of directions along the horizontal and vertical axes ($x/y$ flips) and (ii) the sample entropy (SamEn) measure of the $x$-axis only (Richman and Moorman, 2000). Note that, in this context, the eventual "complexity in response trajectories may be taken as evidence for a formal dynamical process at work" (Hehman, Stolier, and Freeman, 2014, p.9).

### 3.2.3 Some critical issues

DGA is a simple, fast, and practical tool to extract information from mouse-tracking based experiments. However, despite of this, it can suffer from serious limitations both technically and theoretically. In what follows, we outline some of the most important critical issues that could significantly affect the sensitivity and strength of DGA-based studies.
The first critical aspect concerns the process of *time-normalization* which is applied on the raw mouse trajectories. Although such a procedure would guarantee averaging and comparisons across multiple trials, it could result in a distortion of the original movement profiles with the consequence of a significant loss of information. This problem is widely known in the signal processing literature where researchers have dealt with problems referring to signals alignment, data stretching, data normalization and wrapping, and have proposed several techniques, more sophisticated than the simple linear time-normalization, that try to model the empirical noisy time-based profile by respecting its relevant characteristics (e.g., see Lucero and Koenig, 2000; Ramsay, 2006; Tang and Müller, 2008). A second relevant issue concerns the *averaging procedure*. Although it would simplify the subsequent DGA analyses, it could possibly lead to inappropriate results especially when empirical profiles of movement show different amplitudes, phase variations, and curvatures (Ramsay, 2006). Moreover, even when profiles would show same directions and curvatures, in some

cases averaging could distort the original profile information (Omar et al., 1999; Matthews et al., 1990). Therefore, the choice of methods to use for averaging (e.g., arithmetic, geometric, block-based) becomes important in order to preserve the original profile information and avoid distortions (e.g., see Brown and Heathcote, 2003). Further, also the so-called *spatial attraction analysis* might be seriously biased. In particular, the MD measure can potentially suffer from the problem of "multiple maxima" in the raw movement curve. Indeed, due to noise that usually affects reaching movements, in several situations mouse trajectories may not completely show clear parabolic patterns with just one inflection point. Moreover, MD can be excessively sensitive to traditional noisy patterns and meaningless for complex sinusoidal, quasi-cyclic, and/or quasi-random movements. In this context, also AUC is not immune from drawbacks. For instance, it can suffer from "compensatory phenomena" depending on the shape of the movement trajectories and can be worthless when movements are complex and/or located in the opposite direction of the distractor (e.g., movements allocated on the right-side of the ideal strength line represented in Fig. 3.1-B/C produce negative area under the curve) (Fekedulegn et al., 2007). Finally, considering the *velocity-acceleration* analysis, the free-model approach used in computing velocity and acceleration yields noisy estimation of the derivatives with the consequence that noise becomes more prominent with increasing of the derivatives (Liu, 2002; Wu et al., 2006).
By and large, as DGA directly handle with the raw trajectories collected over participants and trials, all the aforementioned limitations can be related to the "lack of curve modeling" that affects this approach. Although in some circumstances such an approach can provide modest results (e.g., when analyses deal with few trajectory points), it should be noticed how it completely ignores the "topological data structure" underlying the empirical movement trajectories (Everitt and Pickles, 2004). As a consequence, noise, imperfections, and natural movement fluctuations (e.g., turbulence effects), can deteriorate the quality of the analyses and measures used to summing up the empirical mouse movements (Guo, 2004; Kumaresan,

Tufts, and Scharf, 1984).

## 3.3 Information theoretic approach to spatial information

In this section we describe our approach, named InTMT, in modeling and analysing mouse trajectories from mouse-tracking experiments.

### 3.3.1 General context and motivation

InTMT is an information-theoretic based model that focuses on the *spatial information* of mouse trajectories recorded by means of mouse-tracker methodology.[1] In our approach, the movement paths are firstly modeled by means of an adaptive cumulative function whereas the main features of the movement are subsequently extracted adopting a dedicated entropy decomposition. Basically, InTMT extends the main findings of DGA in some respects, although it provides a new formal framework to model the *motor pauses* usually presented in movement patterns. Indeed, unlike DGA, our approach highlights the important role played by pauses in revealing information about underlying cognitive processes (e.g., decisional conflicts in categorization, decision making under uncertainty) (Plamondon, 1995; Plamondon and Alimi, 1997). Although motor pauses can reflect the mechanical dynamics of the neuromuscular system (Elliott, Helsen, and Chua, 2001), they usually arise from discontinuities in velocity and acceleration of intermittent movement profiles (Meyer et al., 1988; Hwang et al., 2005; Wel et al., 2009). Moreover, as for eye-tracking based paradigms, motor pauses might also reveal important aspects of dynamic attentional processes related to decision and choice making (Johnson et

[1] Over the years, several scholars have tried to apply information theory in psychology. Although such an application has found criticisms (for a review, see Shannon, 1956; Luce, 2003; Laming, 2010; Corning and Kline, 1998), there have been numerous attempts to use its concepts and ideas both theoretically and technically. Recently, for instance, information theory has been applied in modeling the cognitive load in sentence comprehension (Frank, 2013), in measuring how the cognitive system creates new representational structures (Stephen, Dixon, and Isenhower, 2009) and how it solves goals conflicts (Hirsh, Mar, and Peterson, 2012), in quantifying behavioral and neural processes (Rosso, 2007; Besserve et al., 2010), in modeling reaction times (Prado Martín, 2011). Not least, it has been also applied in quantifying some relevant physical as well as distributional characteristics of aiming and reaching movements (e.g., see Seow, 2005; Hoffmann, 2013; Lai, Mayer-Kress, and Newell, 2006; Lai et al., 2005).

al., 2012; Glaholt and Reingold, 2011). Note that, InTMT does not provide a cognitive modelisation of the processes underlying mouse-tracking experiments. On the contrary, it defines a general procedure to extract information from movement trajectories according to a specific model. As a consequence, information theory together with its measures are only used in a descriptive fashion (Baird, 1984).

### 3.3.2 InTMT approach

#### Data modeling

Let $\mathbf{p}_{ij} = (\mathbf{x}_{ij}, \mathbf{y}_{ij}) \in \mathbb{R}^{n_{ij} \times 2}$ be a $n_{ij} \times 2$ array of Cartesian coordinates associated to the streaming of the *x*-*y* data of the computer mouse movements with $n_{ij}$ being the recorded number of movement points varying across subjects ($i = 1...I$) and trials ($j = 1...J$). The ITA-MT adopts the following two-step procedure in modeling the spatial movement data (see Fig. 3.2). *First modeling step.* It represents the raw movement points $\mathbf{p}_{ij}$ by means of a simple and flexible model $\mathbf{m}_{ij}$ which is able to capture the main movement features such as direction, amplitude (i.e., attraction toward the competitive target), and frictions (pauses). To do this, we first remove eventual imprecision due to hand motor controls and/or computer mouse adjustments. In particular, the *x*-*y* coordinates in $\mathbf{p}_{ij}$ that are located near to the starting point (symbol "0" in Fig. 3.2-A) and the points beyond the borders of targets (symbols "C1" and "C2" in Fig. 3.2-A), are all removed by applying a predefined filter which defines the area for acceptable *x*-*y* coordinates. Next, the refined $\mathbf{p}_{ij}$ is transformed into a $n_{ij} \times 1$ vector $\mathbf{p}^c_{ij} \in \mathbb{R}^+_{\{0\}}$ of angles (in radians) by means of any arctangent based function (e.g., *atan2*). Lastly, the angles in $\mathbf{p}^c_{ij}$ are modeled by means of a linear histogram whose support varies from $[0, \pi/2]$ (Fisher, 1995). Note that, the histogram of movement points (Fig. 3.2-C) represents the directions of the *x*-*y* movements naturally expressed in $\mathbb{R}^{n_{ij} \times 2}$ as points on the positive real line $\mathbb{R}^+_{\{0\}}$ whereas the number of points for each bin of the histogram represent the eventual frictions during the movement. Therefore, the horizontal

dimension of the histogram reproduces the direction and the amplitude of the movement whereas the vertical dimension quantifies the motor pauses that are codified in the Cartesian plane.[2]

*Second modeling step.* It transforms the measures in $\mathbf{p}_{ij}^{c}$ by means of a smooth cumulative function $\mathbf{m}_{ij}$. In particular, $\mathbf{m}_{ij}$ can be obtained by any smoothing procedure, for instance by non-parametric Kernel estimation:

$$\mathbf{m}_{ij} = (n_{ij})^{-1} \sum_{t=1}^{n_{ij}} H\left(\frac{\mathbf{p}_{ij}^{c} - p_{ijt}^{c}}{h}\right)$$

where $H(\mathbf{p}_{ij}^{c})$ is the Kernel function that is usually numerically estimated whereas $h$ is the so-called bandwidth parameter that can be estimated, for instance, via cross-validation by minimizing the mean integrated squared error (MISE) (e.g., see Azzalini and Bowman, 1997; Rio and Estevez-Perez, 2012).

Note that as a consequence of the properties of the cumulative functions (Burr, 1942), $\mathbf{m}_{ij}$ results to be more robust and regular in case of noisy and/or irregular movement profiles. This means, for instance, that fast directional changes in $\mathbf{p}_{ij}$ result in slower changes in $\mathbf{m}_{ij}$ and allows the model to focus expressly on the relevant movement events occurred in the reaching task. Interestingly, some structural relations exist between histograms of movement points and cumulative functions (see Fig. 3.3). In particular, Fig. 3.3-A represents an histogram of movements where the majority of motor pauses occur toward the cue C1. This aspect is reflected in the height of the cumulative function that results to be less then the threshold line (Fig. 3.3-D, solid curve). By contrast, Fig. 3.3-B shows a typical bimodal histogram associated to movement patterns where motor pauses are (almost) equally distributed toward both the cues C1 and C2 of the movement space. In this situation the cumulative function has

[2]Technically, motor pauses are codified as points showing the same (or almost the same, in a neighborhood) *x*-*y* coordinates.

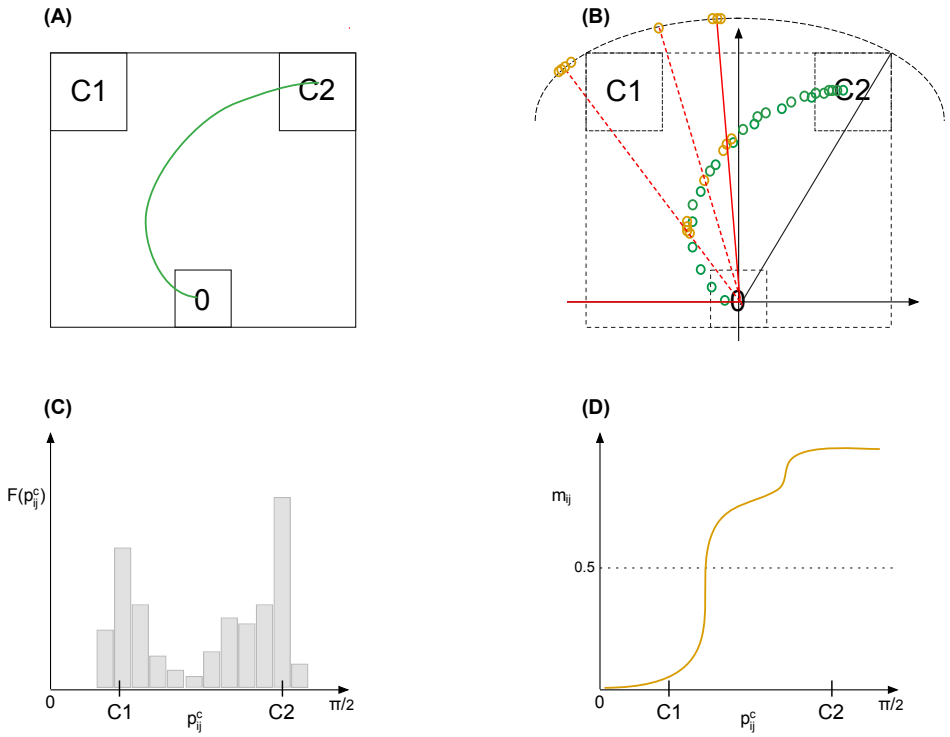



FIGURE 3.2: Conceptual diagram of ITA modeling. (A) Cartesian mouse trajectories $\mathbf{p}_{ij}$ in 2-dimensional plane. (B) *atan2* function transforming $\mathbf{p}_{ij}$ in radians $\mathbf{p}^c_{ij}$. (C) histogram modeling $\mathbf{p}^c_{ij}$. (D) Empirical cumulative distribution function $\mathbf{m}_{ij}$ associated to the histogram of movement points

height closed to the threshold line (Fig. 3.3-E). Finally, Fig. 3.3-C represents movement patterns with motor pauses toward the cue C2. As a consequence the cumulative function shows height greater than the threshold line (Fig. 3.3-F). Moreover, such relations became meaningful in the context of mouse-tracking experiments. In particular, panel A in Fig. 3.3 represents the typical case in which the cue C1 produces attraction as well as motor pause at the same time. Indeed, the support of the histogram spans (in randians) from C1 to C2 whereas the majority of frequencies is located on C1. On the contrary, panel C in Fig. 3.3 reflects the situation in which mouse movements show larger motor pauses toward the final selected cue C2 while showing lower frictions toward C1 even though they span all the interval ranging from C1 to C2. This situation might occur when fast and hasty mouse movements do not relate with a real attraction of the competitive cue C1. Instead, panel B in Fig. 3.3 codifies the situation where

both C1 and C2 cues exercise attraction on mouse movements by generating the situation of maximum uncertainty. Note that our model allows to represent all the intermediate situations between movement amplitude and motor pause, as they are codified by two different characteristics of the histogram. Consequently, our approach becomes flexible also in modeling situations where motor pauses and amplitude of the movements are unrelated.

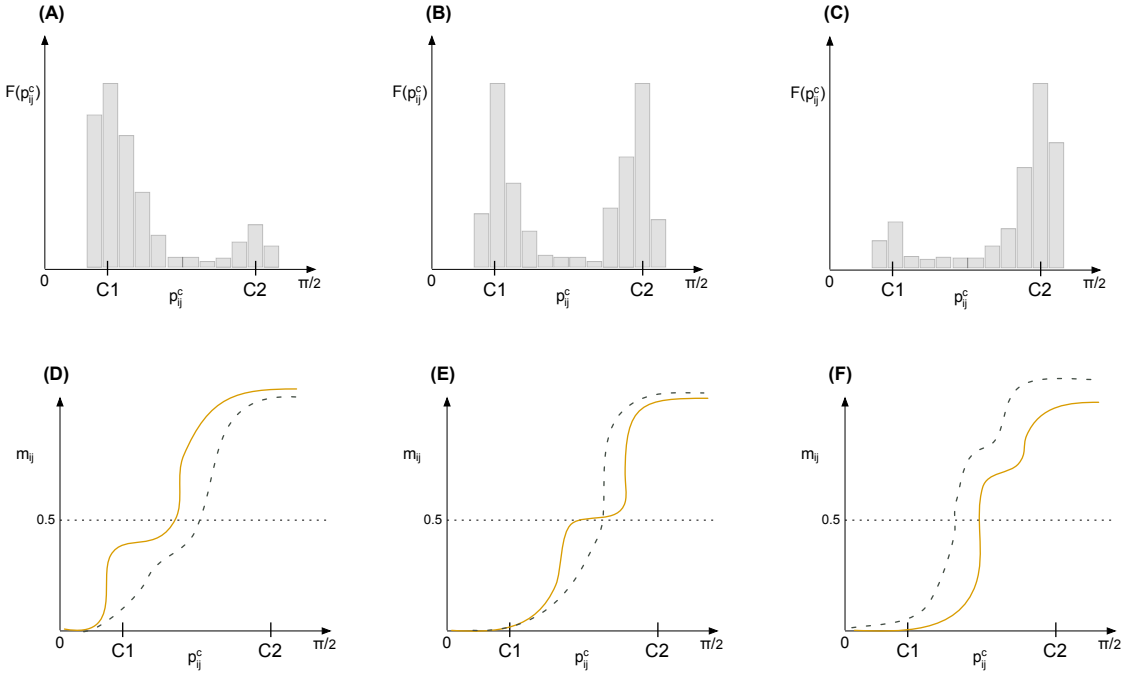



FIGURE 3.3: Conceptual relations between histograms of movement points and cumulative functions (in yellow). (A-D) Patterns with ideal attraction toward C2. (B-E) Patterns with attraction equally distributed between C2 and C1. (C-F) Patterns with no attraction toward C2. Note that dotted gray curves represents the patterns of pure movement (without pauses) whereas the value of 0.5 in the horizontal axis represents the threshold line.

## Data analysis

While DGA directly extracts movement features from the raw movement profiles, in the InTMT approach all the information characterizing the movement are coded with the cumulative functions $\mathbf{m}_{ij}$'s that are further manipulated by means of a dedicated entropy calculus. More formally, the

model $\mathbf{m}_{ij}$ is mainly synthesized by the *cumulative residual entropy* (CRE) measure:

$$\psi(\mathbf{m}_{ij}) = (1 - \mathbf{m}_{ij})^T \log (1 - \mathbf{m}_{ij}) \tag{3.1}$$

that expresses the information contained in the cumulative function of the movement.[3] Note that $\psi(\mathbf{m}_{ij})$ can be directly compared with MD (e.g., it tends to be large with increasing the amplitude of the movement toward the unselected cue C1). However, because it is calculated on the model $\mathbf{m}_{ij}$ rather than on $\mathbf{p}_{ij}$, unlike MD, the measure $\psi(\mathbf{m}_{ij})$ appears more robust to noisy and irregular movement paths. Further, the entropy representation allows to independently quantify the contribution of movements and motor pauses as follows:

$$\psi(\mathbf{m}_{ij}) = \xi(\mathbf{m}_{ij}) + \zeta(\mathbf{m}_{ij}) \tag{3.2}$$

where $\xi(\mathbf{m}_{ij})$ is the cumulative residual entropy of the pure movement (i.e., movement without motor pauses) whereas $\zeta(\mathbf{m}_{ij})$ is the corresponding entropy associated to the pauses. This is an interesting results, especially since it allows a simple decomposition of the motor information provided by the mouse trajectories. In addition, the equality 3.2 can be further expanded as follows:

$$\psi(\mathbf{m}_{ij}) = \xi(\mathbf{m}_{ij}) + \zeta_1(\mathbf{m}_{ij}) + \zeta_2(\mathbf{m}_{ij}) \tag{3.3}$$

where $\zeta_1(\mathbf{m}_{ij})$ and $\zeta_2(\mathbf{m}_{ij})$ are the entropies of the motor pauses occurred during the movements toward the correct cue (C1) or the competitive cue (C2), respectively (see also Fig. 3.2-A). In this way, the entropy calculus

[3]This measure has been introduced by Rao et al. (Rao et al., 2004) by generalizing the main properties of Shannon entropy. As a number of studies revealed, CRE shows general and nice mathematical properties(Wang et al., 2003; Baratpour and Rad, 2012; Di Crescenzo and Longobardi, 2009). Particularly, (i) its definition is valid for both continuous and discrete cases, (ii) it can be easily computed from the sample data as its estimation converges in the limit to the true value, and (iii) can be also computed for bimodal distributions. On this basis, also other traditional information-theoretic measures have been extended to deal with CRE, such as the cumulative residual KL measure (Park, Rao, and Shin, 2012), dynamic cumulative residual entropy (Navarro, Aguila, and Asadi, 2010), and cumulative mutual information (Wang and Vemuri, 2007).

provides insights into the motor paths recorded during mouse-tracking tasks.

At this point, we need to quantify all components of the right side of Eq. 3.3. To this aim, we resort to use the following procedure that is in line with the descriptive approach underlying InTMT.

First, we compute an approximation of the *zeroth-frequency moment* of $\mathbf{p}_{ij}^{c}$ as

$$\mathbf{p}_{ij}^{0} = F_0(\tilde{\mathbf{p}}_{ij}^{c})$$

where $\tilde{\mathbf{p}}_{ij}^{c}$ is the nearest integer vector to $\mathbf{p}_{ij}^{c}$ whereas $F_0$ is implemented by any appropriate algorithm (Hall Jr, 1948; Coppersmith and Kumar, 2004). The new vector $\mathbf{p}_{ij}^{0}$ contains only distinct elements of $\mathbf{p}_{ij}^{c}$ and refers to the vector of *pure movement points*.

Second, the cumulative model of the pure movement is computed as previously described, namely:

$$\mathbf{m}_{ij}^{0} = (n_{ij})^{-1} \sum_{t=1}^{n_{ij}} H\left(\frac{\mathbf{p}_{ij}^{0} - p_{ijt}^{0}}{h}\right)$$

where $H(\mathbf{p}_{ij}^{0})$ is the Kernel numerically estimated whereas $h$ refers to the bandwidth parameter.

Third, the entropy measures associated with the motor pauses are obtained as:

$$\zeta(\mathbf{m}_{ij}) = \left| \, \psi(\mathbf{m}_{ij}) + (1 - \mathbf{m}_{ij}^{0})^{T} \log\left(1 - \mathbf{m}_{ij}^{0}\right) \right|$$

$$\zeta_1(\mathbf{m}_{ij}) = \left| \, \psi(\mathbf{m}_{ij}) + (1 - \boldsymbol{\alpha}_{ij}^{+})^{T} \log\left(1 - \boldsymbol{\alpha}_{ij}^{+}\right) \right|$$

$$\zeta_2(\mathbf{m}_{ij}) = \left| \, \psi(\mathbf{m}_{ij}) + (1 - \boldsymbol{\alpha}_{ij}^{-})^{T} \log\left(1 - \boldsymbol{\alpha}_{ij}^{-}\right) \right|$$

where:

$$\boldsymbol{\alpha}_{ij}^{+} = \{m_{ijt} \,|\, m_{ijt}^{0} \geq m_{ijt}\} \quad \text{and} \quad \boldsymbol{\alpha}_{ij}^{-} = \{m_{ijt} \,|\, m_{ijt}^{0} < m_{ijt}\} \qquad (t = 1 \ldots n_{ij})$$

Finally, the quantity $\xi(\mathbf{m}_{ij})$ follows by the definition 3.2:

$$\xi(\mathbf{m}_{ij}) = \psi(\mathbf{m}_{ij}) - \zeta(\mathbf{m}_{ij})$$

Note that also in this context we can define a taxonomy of relationships between the models $\mathbf{m}_{ij}$ and $\mathbf{m}^0_{ij}$ (see Fig. 3.3). In particular, when mouse trajectories and motor pauses are in the direction of the attractor cue C1 (Fig. 3.3-A/D), the curve $\mathbf{m}_{ij}$ (yellow solid line) always dominates $\mathbf{m}^0_{ij}$ (dotted black line). On the contrary, when movements and pauses are toward the correct cue C2 (Fig. 3.3-C/F), $\mathbf{m}^0_{ij}$ always dominates $\mathbf{m}_{ij}$. Instead, in the case of maximum uncertainty (Fig. 3.3-C/E) both the models are interlaced. It should be finally noted that these properties play an important role in defining the components $\zeta_1$ and $\zeta_2$ as parts of $\zeta(\mathbf{m}_{ij})$.[4]

An additional issue concerns how two or more movement profiles can be compared. InTMT provides a direct way to compare profiles in terms of entropy. In particular, given two movements models $\mathbf{m}^a_{ij}$ and $\mathbf{m}^b_{ij}$ one can compute the *profile overlap* in terms of distance by means of the *cumulative KL divergence* (Baratpour and Rad, 2012; Park, Rao, and Shin, 2012):

$$\chi(\mathbf{m}^a_{ij}, \mathbf{m}^b_{ij}) = \sum_{t=1}^{n_{ij}} (1 - m^a_{ijt}) \left[ \frac{1 - m^a_{ijt}}{1 - m^b_{ijt}} - \log\left( \frac{1 - m^a_{ijt}}{1 - m^b_{ijt}} \right) - 1 \right] \tag{3.4}$$

where $\chi(\mathbf{m}^a_{ij}, \mathbf{m}^b_{ij}) \in [0, \infty]$ and equals to zero when $\mathbf{m}^a_{ij} = \mathbf{m}^b_{ij}$.

## 3.4 Some exemplary applications

In this section we will shortly present two applications of the InTMT approach. In particularly, we will first describe the main characteristics of the proposed approach by performing a trial-by-trial analysis on the data coming from a standard categorization task (Dale, Hindy, and Spivey,

[4] Note that the identity $\zeta(\mathbf{m}_{ij}) = \zeta_1(\mathbf{m}_{ij}) + \zeta_2(\mathbf{m}_{ij})$ holds.

2006). Lastly, we will show how to apply the InTMT procedure on a cross-sectional analysis of lexical decision data (Barca and Pezzulo, 2012).

*Application 1*. In this first application we will perform a simple trial-by-trial analysis on a proper subset of data that are already published (Dale, Hindy, and Spivey, 2006). This is a simple lexical categorization task where basic level animal names (e.g., *gorilla*) were assigned to their respective superordinate category (e.g., *mammal*). Note that the original experiment was executed using the mouse tracking methodology. Particularly, they were asked to categorize lexical stimuli in the correct category by choosing between two possible assignments (e.g., *gorilla*: *mammal* vs. *reptile*). The underlying hypothesis is that stimuli showing higher proximity with the incorrect category will result in larger mouse trajectories (i.e., the higher uncertainty in categorizing the stimulus, the larger *x*-*y* trajectory). The sample is represented by 31 right-handed students from the Cornell University. Lexical stimuli (125 names of animals) belonging to the natural superordinate categories of mammal, fish, reptile, bird, and insect, were extracted from the McRae repository . Experiment was run using RealBasic whereas the computer-mouse sample rate is approximately 40Hz. Figure 3.4-3.7 show four empirical patterns who are characterized by some nice properties concerning movements, pauses, and changes in directions. In particular, Figure 3.4 shows a patter of movement with no attraction toward the competing category *reptile* while movement points are all located on the area of the correct category *mammal*. As explained from the standard ballistic hypothesis in reaching movements, such a patter contains two main pauses at the beginning and at the end of the movement. As a consequence, they would reflect planning motor executions and their finalisation, respectively. Therefore, the movement appears straight and clean providing a global entropy equal to $\psi(\mathbf{m}) = 9.79$ which is almost entirely characterized by the movement component $\xi(\mathbf{m}) = 9.36$ (the quantity $\zeta_2(\mathbf{m}) = 0.43$ refers to a short pause which happened after the movement started).

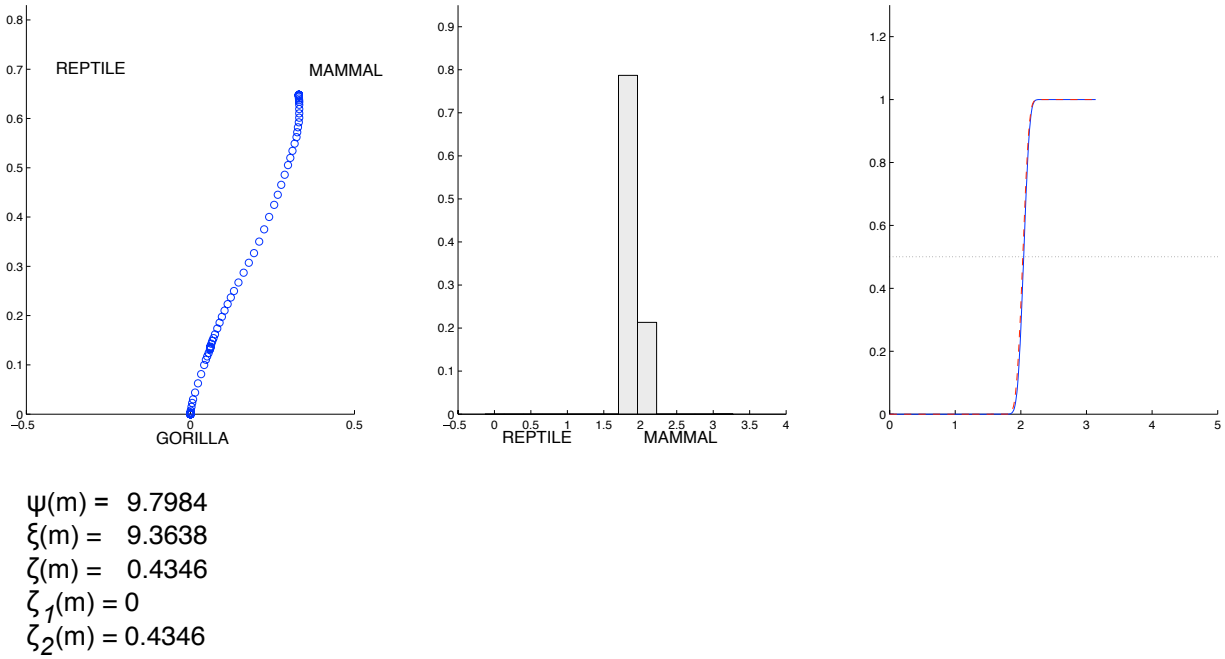
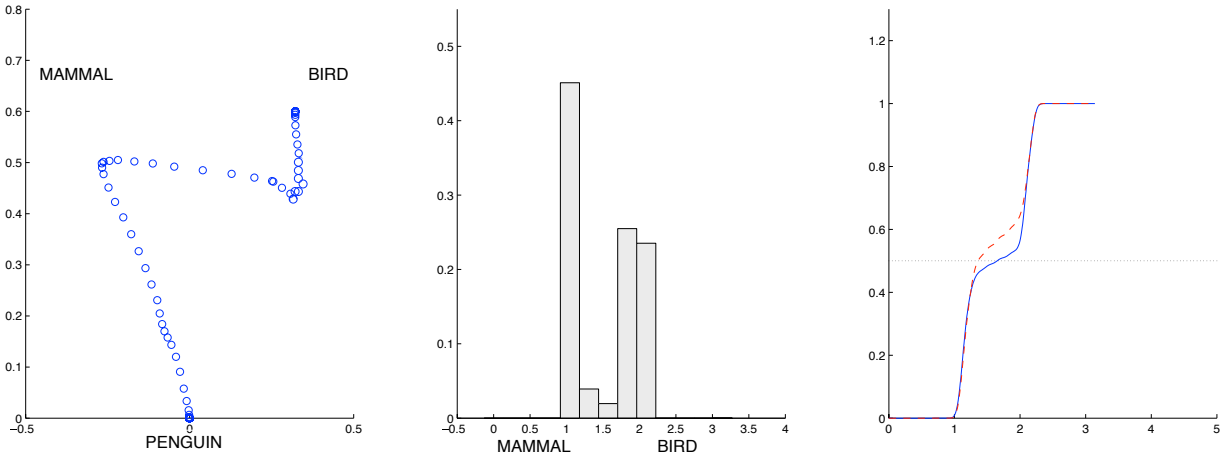


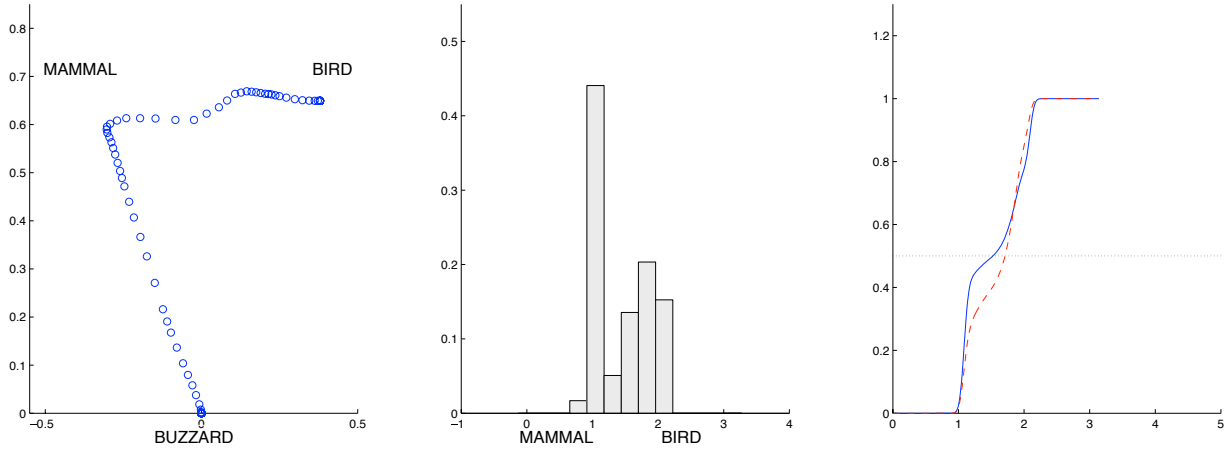

FIGURE 3.4: Application 1: subject 10, trial 61. Note that, $\zeta_1(\mathbf{m})$ refers to the competing category 'reptile' whereas $\zeta_2(\mathbf{m})$ to 'mammal'.

Figure 3.5 shows a typical patter of attraction which is also characterized by pauses toward the competing category. Accordingly to the hypothesis underlying such task, participant showed uncertainty in classifying the stimulus *penguin* in the correct category *bird* because of the lexical attraction provided by the other category *mammal*. This is reflected in the points of movement, especially in their location, as well as in the pauses occurring toward the competing category. As a consequence, the global entropy $\psi(\mathbf{m}) = 30.31$ is mainly ascribed to the movement ($\xi(\mathbf{m}) = 23.11$) and the pause occurred toward the opposite category *mammal* ($\zeta_1(\mathbf{m}) = 6.22$).
Similarly, Figure 3.6 can be interpreted as in the previous case although here the pause component ascribed to the correct category *bird* is more pronounced ($\zeta_2(\mathbf{m}) = 2.50$).
Figure 3.7 shows an irregular pattern of movement with something like a *stenosis* in the first part of the movement and a slow shift toward the correct category. Note also that, on the las part of the movement the pattern still shows a stop occurring before clicking the final correct label *mammal*. Although such an irregular pattern cannot be directly classified with




$\psi(m) = 30.3122$
$\xi(m) = 23.1126$
$\zeta(m) = 7.1996$
$\zeta_2(m) = 0.9721$
$\zeta_1(m) = 6.2275$

FIGURE 3.5: Application 1: subject 13, trial 11. Note that, $\zeta_1(\mathbf{m})$ refers to the competing category 'mammal' whereas $\zeta_2(\mathbf{m})$ to 'bird'.




$\psi(m) = 57.1654$
$\xi(m) = 49.0465$
$\zeta(m) = 8.1189$
$\zeta_2(m) = 2.5066$
$\zeta_1(m) = 5.6123$

FIGURE 3.6: Application 1: subject 25, trial 103. Note that, $\zeta_1(\mathbf{m})$ refers to the competing category 'mammal' whereas $\zeta_2(\mathbf{m})$ to 'bird'.

a regular parabolic layout, it still presents a "proper way" to codify the uncertainty which would be completely lost in the DGA approach. The

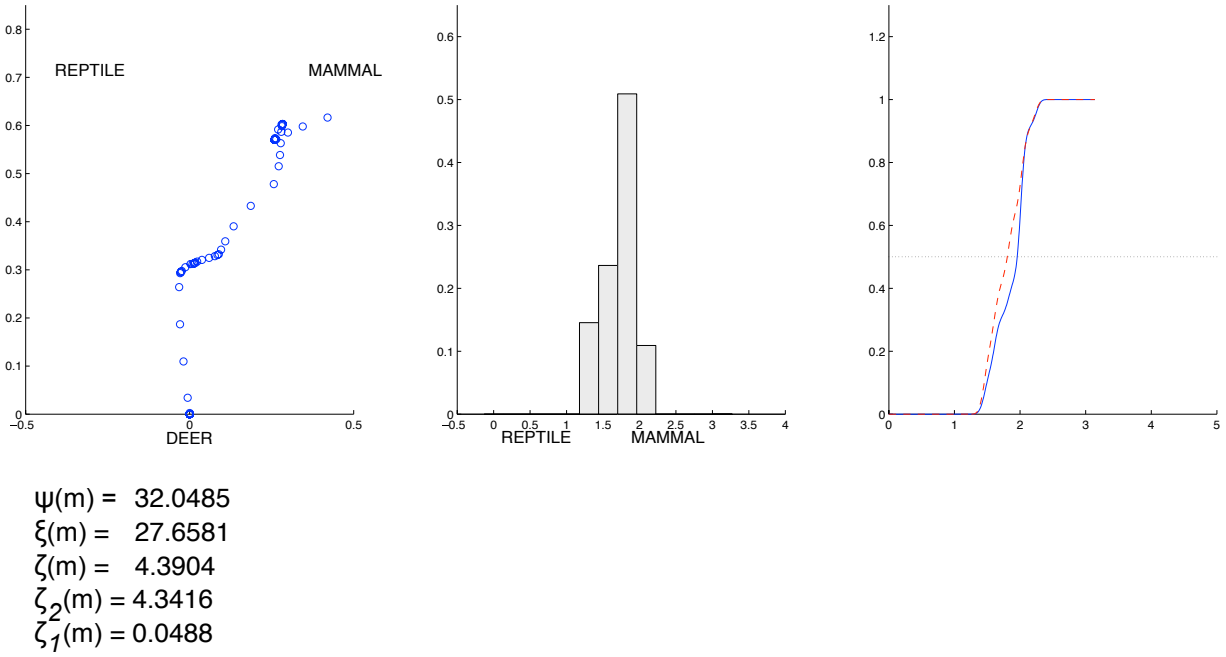



FIGURE 3.7: Application 1: subject 30, trial 87. Note that, $\zeta_1(\mathbf{m})$ refers to the competing category 'reptile' whereas $\zeta_2(\mathbf{m})$ to 'mammal'.

entropy decomposition follows this configuration as expressed by the indices showed in the figure.

*Application 2*. In this second application we will perform a simple statistical analysis considering the InTMT measures for all trials together. We use already published psycholinguistic data where linguistic stimuli (e.g., *latte*, *ghebo*) are categorized as "lexical" or "non-lexical" (Barca and Pezzulo, 2012). The set of stimuli contained 96 lexical as well as non lexical items from the Italian language (stimuli were equally five letters long). These stimuli were manipulated according to two dimensions: frequency (i.e., words with high occurrence vs. words with low occurrence) and type (i.e., pseudowords vs. letters strings). The underlying hypothesis here is that more ambiguous items, such as low frequency words and pseudowords, would show more uncertainty during the categorization task. To do that, 22 right-handed participants with normal/corrected vision and ages ranging from 20 to 35 years, were recruited while kinematic data were recorded with the mouse tracking methodology. With regards to the

aim of our proposal, we have re-analysed data by applying the InTMT procedure. Figure 3.8 shows the entropy decomposition for each of the four experimental levels. As expected, compared with the condition letter strings (LS) the task with pseudowords (PW) generates higher global entropy $\psi(\mathbf{m})$ which is also reflected in higher entropy of movement $\xi(\mathbf{m})$ as well as pauses $\zeta(\mathbf{m})$. In particular, such a condition shows higher pauses toward the competing categories of the stimuli $\zeta_1(\mathbf{m})$ than the corresponding LS condition. Generally, the task under the condition PW seems to generate more difficulties during the categorization task. On the contrary, the conditions low and high frequencies (HF vs LS) seem to be overlapped in the entropy components. In particular, the condition LF is associated with highest global entropy $\psi(\mathbf{m})$. Note that, such a result can be understood under the components $\xi(\mathbf{m})$ and $\zeta_2(\mathbf{m})$, namely the condition LF seems to generate higher entropy of movement as well as pauses toward the correct category C2. However, surprisingly, the condition HF shows slightly higher entropy of pauses $\zeta_1(\mathbf{m})$ toward the competing category C1 then LF does. This counter-intuitive result in the condition HF could be in relation with the occurrence of fast and accidental movements (i.e., words with high frequency can be easily categorized in the correct categories).
In order to complete these preliminary analysis, we ran some tests for differences on the condition HF/LF and LS/PW. Figure 3.9 shows the boxplot with the associated *p-values* computed via Wilcox-test. Note that, we report the condition LS/PW only because the comparison HF/LF showed no significative differences. In line with the findings of the DGA approach (Barca and Pezzulo, 2012), in this particular case all the entropy measures result to be statistically significant.

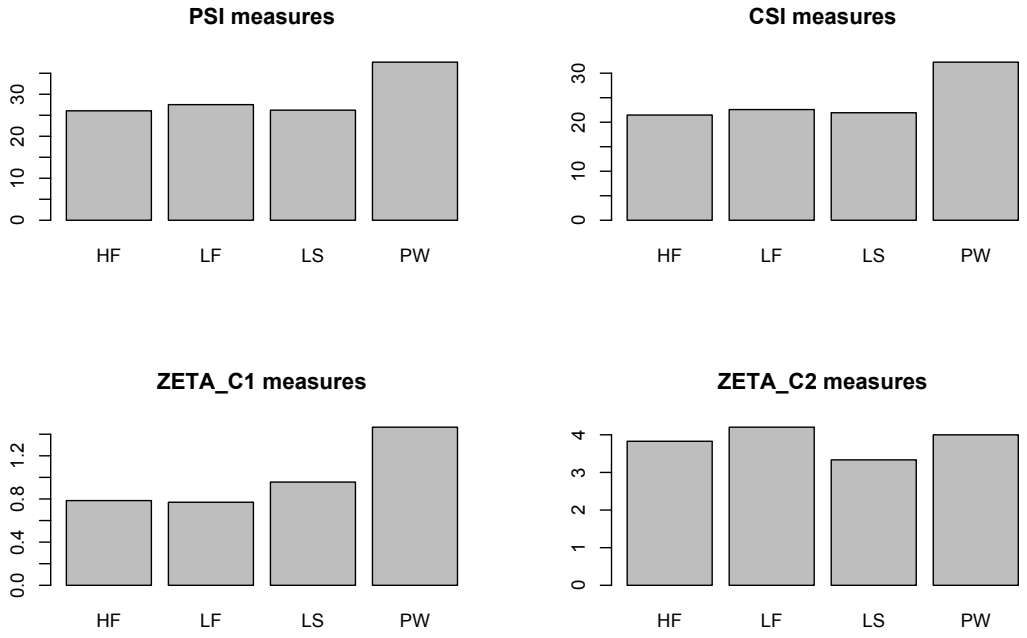



FIGURE 3.8: Application 2: Entropy decomposition for the four experimental levels. Note that, HF and LF mean high and low frequency conditions, whereas LS and PW indicate letter strings and pseudowords conditions, respectively.

## 3.5 Conclusions

In this chapter we described some preliminary findings of our InTMT approach. Unlike the standard DGA approach, it is mainly based on a robust modelization of the mouse trajectories through an adaptive logistic function that contains all the spatial information provided by the empirical *x*-*y* patterns. Moreover, a set of entropy-based measures which include a movement-pause entropy decomposition provided a way to analytically quantify the most relevant spatial information present in the empirical data, namely movements and pauses. In addition, by using suitable distributional distances, InTMT also provided an appropriate framework to compare different movement patterns by means of a proper KL divergence. Finally, in order to show some relevant characteristics of our approach, we conducted two preliminary studies on published datasets. In particular, the first application showed a trial-by-trial analysis where single mouse trajectories were analysed with the InTMT approach whereas

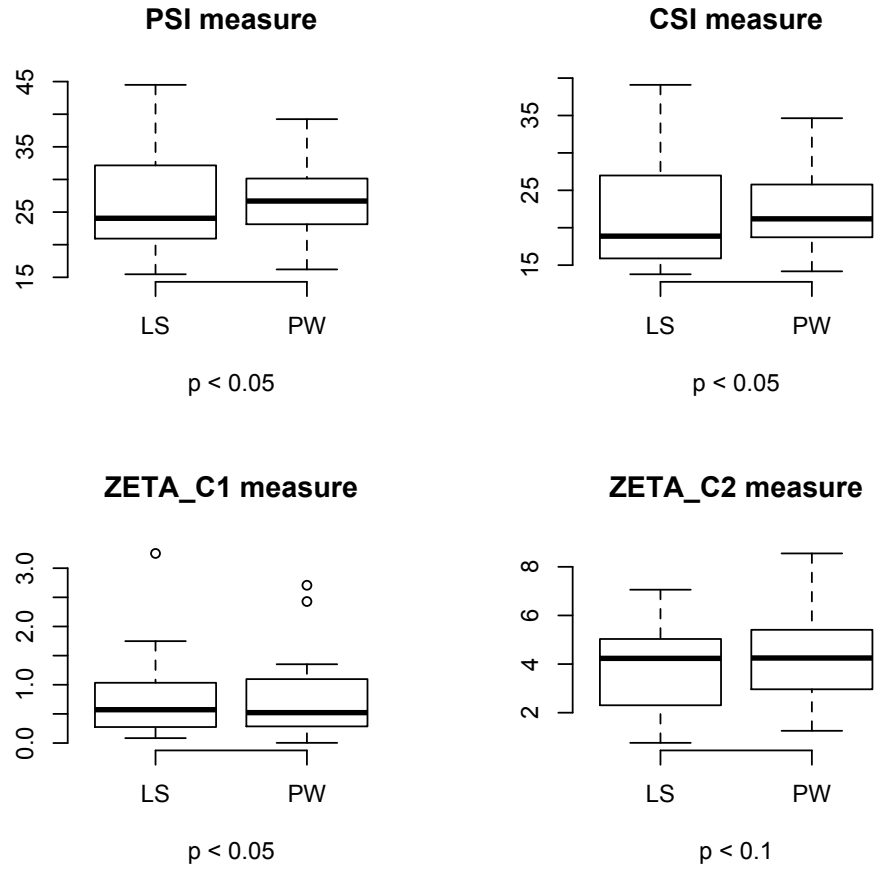



FIGURE 3.9: Application 2: Wilcox-test for the experimental conditions. Note that, HF and LF mean high and low frequency conditions, whereas LS and PW indicate letter strings and pseudowords conditions, respectively.

the second one showed how our entropy decomposition could be a valid alternative in studying the motor components which underly the mouse tracker patters of movement. With particular regards to the second case study, overall our findings confirm what the author already found in their original research although our measures did not show statistical differences in the conditions high frequency (HF) and low frequency (LF) as instead DGA did. In general, this application showed (i) how global differences in terms of movement and/or global trajectories could be better explained by considering their constituent sub-components and (i) how InTMT could be as sensitive as DGA in detecting true effects when these are really present in the empirical data.

Several further venues of research will be taken into account. First, an

extensive and accurate simulation study would test the accuracy, sensitivity, and reliability of the proposed set of measures in detecting the components of the patterns of movement. Therefore, the performance of the measures as well as their rates of credibility would be investigated by comparing with the standard DGA procedure. Lastly, it would be natural to extend our applications to further domains of research (e.g., moral decision making). This would help us in clarifying the role of our measures under controlled and theoretically well-founded scenarios as well as to understand the real portability of our approach. All these extensions are already planned to be performed in the coming months.

# Statistical techniques

# Chapter 4

# Non-convex fuzzy data and fuzzy statistics. A first descriptive approach to data analysis

The content of the chapter has been previously published as: Calcagnì, A., Lombardi, L., & Pascali, E. (2014). Non-convex fuzzy data and fuzzy statistics: a first descriptive approach to data analysis. *Soft Computing*, 18(8), 1575-1588.

## 4.1 Introduction

Fuzzy statistics is a branch of statistical theory devoted to handle with fuzzy data characterized by a particular type of uncertainty, called fuzziness. Nowadays, several statistical models and techniques are available based on different approaches (Buckley, 2004; Coppi, Gil, and Kiers, 2006; Nguyen and Wu, 2006; Taheri, 2003; Viertl, 2006). In particular, some of these models have been defined using the notion of *LR-fuzzy number* (Dubois et al., 1988) that can be considered one of the most important concepts in Fuzzy Set Theory (FST). Moreover, LR-fuzzy numbers have been proven to provide an elegant and compact way to describe a large variety of fuzzy data in different applicative contexts (Dubois and Prade, 2000; Ross, 2009; Verkuilen and Smithson, 2006). One of the most important features of LR fuzzy numbers is that they are characterized by convex shapes. However, one potential limitation of convex representations is that, in general, they might not be flexible enough to describe complex structures in

the data. For example, in contexts like human ratings and fuzzy decision making, observed data can easily show some relevant level of non-convexity. Within an engineering perspective, the non-convexity problem could be straightforwardly reconverted into the standard convex representation by using ad-hoc data transformation procedures (e.g., the Graham Scan algorithm). However, this approach may be troublesome when data are genuinely characterized by non-convex features as the transformation can distort the original information stored in the data. Unlike this approach, our contribution proposes an alternative perspective, which is instead based on a direct representation of non-convex fuzzy data. Interestingly, the non-convexity problem has already received some attention in the FST literature (Facchinetti and Pacchiarotti, 2006; Garibaldi and John, 2003; Garibaldi et al., 2004; Lee, Kim, and Jang, 2008; Reuter, 2008; Viertl, 1996). However, the development of specific LR-fuzzy numbers based statistical techniques to analyse such type of data, as far as we know, has not been proposed yet. To fill this gap, in this article we introduce a new computational definition for non-convex fuzzy numbers which extends the traditional definition of *LR-fuzzy numbers*. Moreover, we also present a novel regression model for crisp input and non-convex fuzzy output which generalizes previous approaches based on the least squares framework for fuzzy linear regression (Celmiņš, 1987; Coppi et al., 2006; Diamond, 1988; D'Urso, 2003; Kacprzyk and Fedrizzi, 1992).

The reminder of the chapter is structured as follows. The second section is devoted to describe the concept of non-convex fuzzy numbers whereas the third section exposes the fuzzy regression model together with its main properties. To better highlight the importance of non-convexity in empirical data, the fourth section illustrates five applications of the fuzzy regression model to some behavioural and socioeconomic data collected using different procedures (e.g., fuzzy scales of measurement and fuzzy measurement systems) as well as previously published datasets. Finally, the fifth section concludes this chapter by providing final comments and

suggestions about future extensions of this work.

## 4.2 Non-convex fuzzy numbers: a computational proposal

We briefly recall some basic definitions about fuzzy sets and LR fuzzy numbers.

### 4.2.1 Convex fuzzy sets and LR fuzzy numbers

In general, a fuzzy set $\tilde{A}$ can be described by its continuous family of nested subsets (called $\alpha$-sets):

$$\tilde{A}_\alpha = \{x \in U \mid \mu_{\tilde{A}}(x) > \alpha\} \qquad \text{with } \alpha \in ]0,1] \tag{4.1}$$

with $U$ and $\mu_{\tilde{A}}$ being the universal set and the membership function of $\tilde{A}$, respectively. A *convex* fuzzy set is a fuzzy set $\tilde{A}$ satisfying the following condition:

$$\forall\, a,b,c \in \tilde{A}_0 \mid a \leq b \leq c: \;\; \mu_{\tilde{A}}(b) \geq \; min[\mu_{\tilde{A}}(a), \mu_{\tilde{A}}(c)] \tag{4.2}$$

with $\tilde{A}_0 = \{x \in U \mid \mu_{\tilde{A}}(x) > 0\}$ being the *support* of $\tilde{A}$. Similarly, we say that $\tilde{A}$ is convex if its $\alpha$-sets are all convex sets. The *core* of a fuzzy set $\tilde{A}$ is the collection $\tilde{A}_g = \{x \in U \mid \mu_{\tilde{A}}(x) = max_{y \in U}\; \mu_{\tilde{A}}(y)\}$ of all its maximal points. The *height* of $\tilde{A}$ is defined as $hgt(\tilde{A}) = max\;[\mu_{\tilde{A}}(x)]$. More precisely, if $hgt(\tilde{A}) = 1$ (resp. $0 < hgt(\tilde{A}) < 1$), the fuzzy set $\tilde{A}$ is called *normal* (resp. *subnormal*). Notably, if $\tilde{A}$ is a normal convex fuzzy set, then for each $\alpha \in ]0,1[$ the corresponding $\alpha$-set always yields a compact interval with proper minimum ($min$) and maximum ($max$). Moreover, if $\tilde{A}$ satisfies the following three conditions:

c.1 normality

c.2 convexity

c.3 unimodality (single core representation for $\tilde{A}$)

then it is called a LR-fuzzy number, denoted with $\tilde{a}$ (Dubois et al., 1988; Hanss, 2005; Viertl, 2011). Finally, if we also consider two monotonic decreasing and left-continuous functions (called *shape functions*) for $\tilde{a}$:

$$L: \mathbb{R}^+ \to [0,1] \quad \text{and} \quad R: \mathbb{R}^+ \to [0,1]$$

with properties:

$$L/R(\upsilon)\begin{cases} = 0 & \text{if } \upsilon = 1 \\ = 1 & \text{if } \upsilon = 0 \\ > 0 & \text{if } \upsilon < 1 \\ < 1 & \text{if } \upsilon > 0 \end{cases} \qquad \forall \upsilon \in \mathbb{R}^+$$

then the LR fuzzy number $\tilde{a}$ can be described using the following membership function:

$$\mu_{\tilde{a}}(x) = \begin{cases} L\left(\frac{m-x}{l}\right) & \text{if } x < m \\ R\left(\frac{x-m}{r}\right) & \text{if } x \geq m \end{cases}$$

where $m$, $l$, $r$ are the core, left and right spreads, obeying $l > 0$ and $r > 0$. For a particular choice of the shape functions, namely $L/R(\upsilon) = \max\{0, 1 - \upsilon\}$, one obtains the example represented in Figure 4.1-b (triangular fuzzy number). The geometrical meaning of the parameters $m$, $l$, $r$ are illustrated in the same figure.[1] Finally, the fuzzy number $\tilde{a}$ can be represented by the triple:

$$\tilde{a} = (m, l, r)_{LR}$$

which conveys the main information about the fuzzy set, namely its precision (by means of its core or modal value $m$) and fuzziness (by means of $l$ and $r$). We stress that the analytic expressions of $L$ and $R$ together with $m$, $l$ and $r$ allow us to exactly represent the fuzzy number $\tilde{a}$.

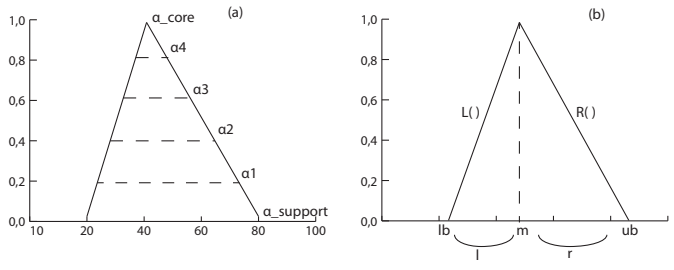



FIGURE 4.1: (a) convex fuzzy set with its $\alpha$-intervals or $\alpha$-sets (dashed lines) (b) convex LR fuzzy number

[1] Note that: $l = (m - lb)$ and $r = (ub - m)$ where $ub$ and $lb$ mean the minimum and maximum of the support, respectively.

### 4.2.2 Non-convex fuzzy sets and non-convex fuzzy numbers

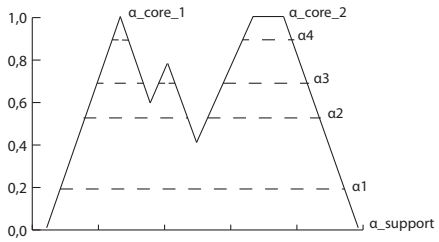



FIGURE 4.2: A graphical representation of a non-convex fuzzy set

Figure 4.2 shows an example of a fuzzy set characterized by two cores with normalized heights. Clearly, the fuzzy set depicted in Figure 4.2 is *non-convex* as, for a suitable choice of $a'$, $b'$, $c' \in U$, the convexity condition is not met. Moreover, we can also notice as $\tilde{A}_{\alpha_3}$ (or rather, $\tilde{A}_{\alpha_2}$, $\tilde{A}_{\alpha_4}$) can be derived by taking the union of disjoint compact intervals, whereas $\tilde{A}_{\alpha_1}$ is a closed and compact $\alpha$-set. Generally speaking, we can say that a fuzzy set is non-convex if some of its $\alpha$-sets represent set theoretical unions of at least two compact disjoint intervals (Viertl, 1996; Viertl, 2011).
We now present a new *operative definition* of non-convex fuzzy numbers which is inspired by the work of (Dubois et al., 1988). More specifically, in this contribution we will limit our attention to a particular but important type of non-convex fuzzy set, called *2-mode fuzzy number* (see Fig. 4.3-a). A 2-mode fuzzy number is a particular kind of fuzzy set which satisfies the following three conditions:

c.4 normality (at least one of the points of its support has full membership value)

c.5 non-convexity

c.6 bimodality (it has no more than two-modes)

About condition (c.5), we restrict the non-convexity property to its simplest representation which corresponds to $\alpha$-sets obtained by taking the union of at maximum two disjoint compact intervals. Finally, notice that conditions (c.5) and (c.6) generalize the standard definition of LR-fuzzy

numbers.[2] We now provide a more formal definition for 2-mode fuzzy numbers. Let $\tilde{B}$ be a fuzzy set. If $\tilde{B}$ satisfies the conditions (c.4-c.6), then $\tilde{b}$ represents a 2-mode fuzzy number. Consider four monotonic decreasing and left-continuous functions:

$$L:\ \mathbb{R}^+ \ \to [0,1] \quad H:\ \mathbb{R}^+ \ \to ]0,1] \quad I:\ \mathbb{R}^+ \ \to ]0,1] \quad R:\ \mathbb{R}^+ \ \to [0,1]$$

with properties:

$$L(\upsilon)\begin{cases} = 0 & \text{if } \upsilon = 1 \\ = t_1 & \text{if } \upsilon = 0 \\ > 0 & \text{if } \upsilon < 1 \\ < t_1 & \text{if } \upsilon > 0 \end{cases} \qquad H(\upsilon)\begin{cases} = t_1 & \text{if } \upsilon = 0 \\ = z & \text{if } \upsilon = 1 \\ > z & \text{if } \upsilon > 0 \\ < t_1 & \text{if } \upsilon < 1 \end{cases}$$

$$\text{with:} \quad \upsilon \in \mathbb{R}^+ \quad t_1 \in ]0,1] \quad \text{and} \quad z < t_1$$

$$I(\upsilon)\begin{cases} = z & \text{if } \upsilon = 1 \\ = t_2 & \text{if } \upsilon = 0 \\ > z & \text{if } \upsilon > 0 \\ < t_2 & \text{if } \upsilon < 1 \end{cases} \qquad R(\upsilon)\begin{cases} = 0 & \text{if } \upsilon = 1 \\ = t_2 & \text{if } \upsilon = 0 \\ > 0 & \text{if } \upsilon < 1 \\ < t_2 & \text{if } \upsilon > 0 \end{cases}$$

$$\text{with:} \quad \upsilon \in \mathbb{R}^+ \quad t_2 \in ]0,1] \quad \text{and} \quad z < t_2$$

By using these functions, the membership function of $\tilde{b}$ can be described in a very general way as follows:

$$\mu_{\tilde{b}}(x) = \begin{cases} L\left(\frac{m_1 - x}{l}\right) & \text{if } x < m_1 \\ H\left(\frac{x - m_1}{h}\right) & \text{if } m_1 < x < m_0 \\ I\left(\frac{m_2 - x}{i}\right) & \text{if } m_0 < x < m_2 \\ R\left(\frac{x - m_2}{r}\right) & \text{if } x > m_2 \end{cases}$$

with $m_1$, $m_2$ and $m_0$ being the modal points and the middle point, respectively; $l$ and $r$ are the external left and right spreads (as defined in the previous section for convex fuzzy numbers); $h$ and $i$ are the *internal spreads*

[2] The name *2-mode fuzzy number* is based on the intuition that fuzzy numbers can be represented by means of the convexity/non-convexity condition. Thus, LR-fuzzy numbers can be named *1-mode fuzzy number* because their $\alpha$-sets are compact and convex sets, whereas *k-modes fuzzy numbers* are fuzzy numbers which $\alpha$-sets are the result of the union of, at maximum, $k$ disjoint components. It is clear that when $k > 1$, the fuzzy numbers are non-convex fuzzy sets.

with $h = (m_0 - m_1)$ and $i = (m_2 - m_0)$, respectively. The following conditions must be satisfied in order to guarantee that $\tilde{b}$ is a two-mode fuzzy number:

c.7 $m_1 < m_0 < m_2$

c.8 $l > 0$

c.9 $r > 0$

c.10 $t_1 < z < t_2$

Moreover, if $L$, $H$, $I$, $R$ are chosen to be linear functions, namely:

$L(\upsilon) = \max\{0, (1 - \upsilon)t_1\}$

$H(\upsilon) = \max\{0, t_1 - \upsilon(t_1 - z)\}$

$I(\upsilon) = \max\{0, t_2 - \upsilon(t_2 - z)\}$

$R(\upsilon) = \max\{0, (1 - \upsilon)t_2\}$

then one obtains the two-mode fuzzy number represented in Figure 4.3-b. Finally, the fuzzy number $\tilde{b}$ can be represented as follows:

$$\tilde{b} = \{(m_0, m_1, m_2, l, r, h, i); (\mu_{m_1}, \mu_{m_0}, \mu_{m_2})\}_{LHIR}$$

where $\mu_{m_1} = t_1$, $\mu_{m_0} = z$, $\mu_{m_2} = t_2$ are the membership values for $m_1$, $m_0$ and $m_2$.

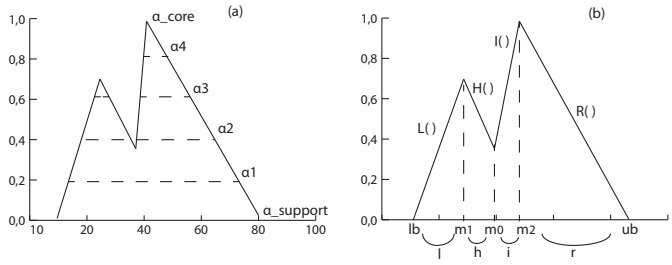



FIGURE 4.3: (a) a non-convex fuzzy set and (b) a LHIR fuzzy number

Under particular conditions, a convex fuzzy number can be understood as a degenerated version of a 2-mode fuzzy number. In particular:

1. when conditions (c.12) is not met the 2-mode fuzzy number degenerates into a *trapezoidal fuzzy number* obeying to the following conditions: $m_1 < m_0 < m_2$, $l > 0$, $r > 0$, $t_1 = z = t_2 = 1$. Moreover, in this particular case, the internal shape functions ($H$ and $I$) assume the following forms $H(\upsilon) = 1$ and $I(\upsilon) = 1$.

2. when conditions (c.7) and (c.10) are not met the 2-mode fuzzy number degenerates into a *triangular fuzzy number* obeying to the following conditions: $m_1 = m_0 = m_2$, $l > 0$, $r > 0$. In this case, the internal shape functions $H(\upsilon)$ and $I(\upsilon)$ (and the corresponding components $\boldsymbol{i}$ and $\boldsymbol{h}$) do not take part in the model representation.

## 4.3 A Fuzzy regression model for 2-mode fuzzy numbers: the *crisp-input/fuzzy-output* case

The relation between crisp independent variables and fuzzy dependent variables plays a particular role in the context of socio-behavioural data. For instance, in cognitive psychology it is common to evaluate the relationships between crisp physical or biological measures (e.g., stimulus brightness, reaction times, eye saccadic movements, levels of blood toxins, etc) and more qualitative psychological responses (e.g., perception of figures, perceived satisfaction, perceived stress, perceived work satisfaction, etc. See: Pashler and Wixted, 2002). The latter variables can be better represented using fuzzy variables. Similarly, in socioeconomic studies, researchers often analyse the relation between crisp (e.g., family income) and fuzzy quantities (e.g., quality of service, quality of teaching, etc) (Benítez, Martín, and Román, 2007; Biswas, 1995; Chan, Kao, and Wu, 1999; Chang and Yeh, 2002; Lalla, Facchinetti, and Mastroleo, 2005). Therefore, it seems worthwhile to develop a model which is able to describe the relations between these different types of variables. Finally, although some empirical contexts may require different representations (e.g., fuzzy-input fuzzy-output relations), in this contribution we wanted to introduce the concept of non-convex data for the most simple case first.

To this end, in this section we propose a least squares fuzzy regression

model for crisp-input/nonconvex fuzzy-output data based on a *generative-type representation*. Unlike more complex models, the generative one provides a good compromise between model flexibility and simplicity.[3] However, to avoid confusion, we stress that the model proposed in this article is only able to deal with 2-mode fuzzy numbers characterized by piecewise linear membership functions (as those represented in Figure 4.3-b) as well as triangular and trapezoidal fuzzy numbers.

### 4.3.1 Model and data analysis

Let $\boldsymbol{X_{n,k}}$ be a $n$ (cases) $\times$ $k$ (variables) matrix representing the set of independent variables. Moreover, let

$$\tilde{\boldsymbol{Y}}_{\mathcal{M}_1} = \{\boldsymbol{m_1}, \boldsymbol{m_0}, \boldsymbol{m_2}, \boldsymbol{l}, \boldsymbol{h}, \boldsymbol{i}, \boldsymbol{r}\} \quad \text{and} \quad \tilde{\boldsymbol{Y}}_{\mathcal{M}_2} = \{\boldsymbol{\mu}_{\boldsymbol{m_1}}, \boldsymbol{\mu}_{\boldsymbol{m_0}}, \boldsymbol{\mu}_{\boldsymbol{m_2}}\}$$

be two fuzzy matrices representing the fuzzy dependent variable. In particular, $\tilde{\boldsymbol{Y}}_{\mathcal{M}_1}$ and $\tilde{\boldsymbol{Y}}_{\mathcal{M}_2}$ denote the data for the horizontal axis and the vertical axis, respectively, in the two dimensional representation of the fuzzy set. Note that the components of these two structures are $n \times 1$ data vectors. Now, we propose a model representation for $\tilde{\boldsymbol{Y}}_{\mathcal{M}_1}$ and $\tilde{\boldsymbol{Y}}_{\mathcal{M}_2}$. In particular, let

$$\tilde{\boldsymbol{Y}}^*_{\mathcal{M}_1} = \begin{cases} \boldsymbol{m_0^*} = \boldsymbol{X a} \\ \quad \boldsymbol{m_1^*} = \boldsymbol{m_0^*} b + \mathbf{1} r_1 \\ \quad\quad \boldsymbol{l^*} = \boldsymbol{m_1^*} c + \mathbf{1} r_2 \\ \quad \boldsymbol{m_2^*} = \boldsymbol{m_0^*} e + \mathbf{1} r_3 \\ \quad\quad \boldsymbol{r^*} = \boldsymbol{m_2^*} d + \mathbf{1} r_4 \end{cases} \qquad \tilde{\boldsymbol{Y}}^*_{\mathcal{M}_2} = \begin{cases} \boldsymbol{\mu}^*_{\boldsymbol{m_0}} = \boldsymbol{X \alpha} \\ \quad \boldsymbol{\mu}^*_{\boldsymbol{m_1}} = \boldsymbol{\mu}^*_{\boldsymbol{m_0}} \beta + \mathbf{1} \xi_1 \\ \quad \boldsymbol{\mu}^*_{\boldsymbol{m_2}} = \boldsymbol{\mu}^*_{\boldsymbol{m_0}} \gamma + \mathbf{1} \xi_2 \end{cases} \tag{4.3}$$

be the two model components, where $\boldsymbol{X}$ is a $n \times (k+1)$ predictor matrix containing in addition also a vector $\mathbf{1}$ of all ones for the intercept, $\boldsymbol{a}$ and $\boldsymbol{\alpha}$ are $(k+1) \times 1$ vectors of regression coefficients, and $b$, $c$, $e$, $d$, $r_1$, $r_2$, $r_3$, $r_4$, $\beta$, $\gamma$, $\xi_1$ and $\xi_2$ are scalars. The model adopts a *generative approach* in which the *first component* $\tilde{\boldsymbol{Y}}_{\mathcal{M}_1}$ represents a hierarchical structure such that the

[3] In the following section we adopt the term *estimation* to indicate the *interpolation procedure* without assuming any inferential meaning (that is to say this approach is based on a descriptive non-inferential rationale).

point $\boldsymbol{m_0^*}$ generates $\boldsymbol{m_1^*}$ and $\boldsymbol{m_2^*}$ and, subsequently, $\boldsymbol{m_1^*}$ and $\boldsymbol{m_2^*}$ generate their external spreads $\boldsymbol{l^*}$ and $\boldsymbol{r^*}$. In a similar way, for the *second component* $\tilde{\boldsymbol{Y}}_{\mathcal{M}_2}$, the membership value $\boldsymbol{\mu^*_{m_0}}$ generates the other two points $\boldsymbol{\mu^*_{m_1}}$ and $\boldsymbol{\mu^*_{m_2}}$. Therefore, the model is able to take into account possible relations between $\boldsymbol{m_0^*}$ and the modes $\boldsymbol{m_1^*}$ and $\boldsymbol{m_2^*}$ as well as the relationships between these points and the spreads. By contrast, the components $\boldsymbol{h^*}$ and $\boldsymbol{i^*}$ can subsequently be derived in a next step by using the already estimated components. Note that both components work in parallel and in an independent fashion. In particular, we do not assume nested or hierarchical relations between the two components. This model representation is in line with a semi-confirmatory approach which assumes that our data are consistent with a generative hypothesis representing possible relations among the modes of the spreads. More precisely, the model captures the dynamic of the spreads as a function of the magnitude of the (estimated) modes. In other words, in some contexts it can be natural to think that the spread (vagueness) in the measure of an empirical phenomenon is to some extent proportional to its intensity (see for example D'Urso and Gastaldi, 2000). Therefore, by considering this assumption, it seems worthwhile to set a descriptive approach based on the minimization of the Euclidean distance between the observed data and the generative model.

### 4.3.2 Parameters estimation: an unconstrained version

In what follows we separately consider the two components of the model.

#### Parameters estimation for $\mathcal{M}_1$

According to the least square criterion the model parameters $\boldsymbol{a}$, $b$, $c$, $e$, $d$, $r_1$, $r_2$, $r_3$, $r_4$ are obtained by minimizing the following distance measure, traditionally used in the fuzzy least squares framework (D'Urso, 2003; D'Urso

and Gastaldi, 2000):

$$\begin{aligned}\mathcal{D}^2 &= \|\boldsymbol{m_0} - \boldsymbol{m_0^*}\|^2 + \|\boldsymbol{m_1} - \boldsymbol{m_1^*}\|^2 + \|\boldsymbol{m_2} - \boldsymbol{m_2^*}\|^2 + \|\boldsymbol{l} - \boldsymbol{l^*}\|^2 + \|\boldsymbol{r} - \boldsymbol{r^*}\|^2 \\ &\equiv \|\boldsymbol{m_0} - \boldsymbol{X}\boldsymbol{a}\|^2 + \|\boldsymbol{m_1} - \boldsymbol{X}\boldsymbol{a}b - \mathbf{1}r_1\|^2 + \|\boldsymbol{m_2} - \boldsymbol{X}\boldsymbol{a}e - \mathbf{1}r_3\|^2 + \\ &+ \|\boldsymbol{l} - \boldsymbol{X}\boldsymbol{a}bc - \mathbf{1}r_1c - \mathbf{1}r_2\|^2 + \|\boldsymbol{r} - \boldsymbol{X}\boldsymbol{a}ed - \mathbf{1}r_3d - \mathbf{1}r_4\|^2\end{aligned} \tag{4.4}$$

In order to solve the minimization problem represented in Eq. 4.4 we equate to zero its partial derivatives (first order conditions), obtaining so the following solutions (by assuming $\boldsymbol{X}^T\boldsymbol{X}$ as non-singular):

$$\begin{aligned}\hat{\boldsymbol{a}} = \frac{1}{k} \cdot (\boldsymbol{X}^T\boldsymbol{X})^{-1}\, \boldsymbol{X}^T[\boldsymbol{m_0} + (\boldsymbol{m_1} - \mathbf{1}r_1)b + (\boldsymbol{m_2} - \mathbf{1}r_3)e + \\ + (\boldsymbol{l} - \mathbf{1}r_1c - \mathbf{1}r_2)bc + (\boldsymbol{r} - \mathbf{1}r_3d - \mathbf{1}r_4)de];\end{aligned} \tag{4.5}$$

with: $k = b^2 + e^2 + b^2c^2 + d^2e^2$

$$\hat{b} = \frac{1}{c^2}\,(\boldsymbol{a}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{a})^{-1}\,\boldsymbol{a}^T\boldsymbol{X}^T[(\boldsymbol{m}_1 - \mathbf{1}r_1) + (\boldsymbol{l} - \mathbf{1}r_1c - \mathbf{1}r_2)c]; \tag{4.6}$$

$$\hat{e} = \frac{1}{d^2}\,(\boldsymbol{a}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{a})^{-1}\,\boldsymbol{a}^T\boldsymbol{X}^T[(\boldsymbol{m}_2 - \mathbf{1}r_3) + (\boldsymbol{r} - \mathbf{1}r_3d - \mathbf{1}r_4)d]; \tag{4.7}$$

$$\hat{c} = \frac{1}{b}\,(\boldsymbol{a}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{a})^{-1}\,\boldsymbol{a}^T\boldsymbol{X}^T(\boldsymbol{l} - \mathbf{1}r_1c - \mathbf{1}r_2); \tag{4.8}$$

$$\hat{d} = \frac{1}{e}\,(\boldsymbol{a}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{a})^{-1}\,\boldsymbol{a}^T\boldsymbol{X}^T(\boldsymbol{r} - \mathbf{1}r_3e - \mathbf{1}r_4); \tag{4.9}$$

$$\hat{r}_1 = \frac{1}{nc^2}\,\mathbf{1}^T[(\boldsymbol{m_1} - \boldsymbol{X}\boldsymbol{a}b) + (\boldsymbol{l} - \boldsymbol{X}\boldsymbol{a}bc - \mathbf{1}r_2)c]; \tag{4.10}$$

$$\hat{r}_3 = \frac{1}{nd^2}\,\mathbf{1}^T[(\boldsymbol{m_2} - \boldsymbol{X}\boldsymbol{a}e) + (\boldsymbol{r} - \boldsymbol{X}\boldsymbol{a}ed - \mathbf{1}r_4)d]; \tag{4.11}$$

$$\hat{r}_2 = \frac{1}{n}\,\mathbf{1}^T(\boldsymbol{l} - \boldsymbol{X}\boldsymbol{a}bc - \mathbf{1}r_1c); \tag{4.12}$$

$$\hat{r}_4 = \frac{1}{n}\,\mathbf{1}^T(\boldsymbol{r} - \boldsymbol{X}\boldsymbol{a}ed - \mathbf{1}r_3d); \tag{4.13}$$

Finally, the above set of equations can be used to compute the best parameter estimates for our model (e.g. using an iterative algorithm).

**Parameters estimation for $\mathcal{M}_2$**

Like for the first component of the model, also for $\tilde{\boldsymbol{Y}}^*_{\mathcal{M}_2}$ the model parameters $\boldsymbol{\alpha}$, $\beta$, $\gamma$, $\xi_1$, $\xi_2$ are obtained by minimizing the following distance measure:

$$\begin{aligned}\mathcal{W}^2 &= \|\boldsymbol{\mu_{m_0}} - \boldsymbol{\mu^*_{m_0}}\|^2 + \|\boldsymbol{\mu_{m_1}} - \boldsymbol{\mu^*_{m_1}}\|^2 + \|\boldsymbol{\mu_{m_2}} - \boldsymbol{\mu^*_{m_2}}\|^2 \\ &\equiv \|\boldsymbol{\mu_{m_0}} - \boldsymbol{X\alpha}\|^2 + \|\boldsymbol{\mu_{m_1}} - \boldsymbol{X\alpha}\beta + \mathbf{1}\xi_1\|^2 + \\ &+ \|\boldsymbol{\mu_{m_2}} - \boldsymbol{X\alpha}\gamma + \mathbf{1}\xi_2\|^2\end{aligned} \tag{4.14}$$

By equating to zero the first order conditions for Eq. 4.14 we obtain the following solutions (by assuming $\boldsymbol{X}^T\boldsymbol{X}$ as non-singular):

$$\begin{aligned}\hat{\boldsymbol{\alpha}} = \frac{1}{(\beta^2+\gamma^2)}\,(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T[\boldsymbol{\mu_{m_0}} + \\ + (\boldsymbol{\mu_{m_1}} - \mathbf{1}\xi_1)\beta + (\boldsymbol{\mu_{m_2}} - \mathbf{1}\xi_2)\gamma];\end{aligned} \tag{4.15}$$

$$\hat{\beta} = (\boldsymbol{\alpha}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\alpha})^{-1}\,\boldsymbol{\alpha}^T\boldsymbol{X}^T(\boldsymbol{\mu_{m_1}} - \mathbf{1}\xi_1); \tag{4.16}$$

$$\hat{\gamma} = (\boldsymbol{\alpha}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\alpha})^{-1}\,\boldsymbol{\alpha}^T\boldsymbol{X}^T(\boldsymbol{\mu_{m_2}} - \mathbf{1}\xi_2); \tag{4.17}$$

$$\hat{\xi_1} = \frac{1}{n}\,\mathbf{1}^T(\boldsymbol{X\alpha}\beta - \boldsymbol{\mu_{m_1}}); \tag{4.18}$$

$$\hat{\xi_2} = \frac{1}{n}\,\mathbf{1}^T(\boldsymbol{X\alpha}\gamma - \boldsymbol{\mu_{m_2}}); \tag{4.19}$$

### 4.3.3 Properties of the model

In this section we only report the most important properties of the model (because they are based on the known principles of least squares approach we omit the proofs from the text).

**Proposition 1.** *Sums of residuals are zero:*

$$\mathbf{1}^T(\boldsymbol{m_0} - \boldsymbol{m_0^*}) = 0 \qquad \mathbf{1}^T(\boldsymbol{m_2} - \boldsymbol{m_2^*}) = 0 \qquad \mathbf{1}^T(\boldsymbol{m_1} - \boldsymbol{m_1^*}) = 0$$
$$\mathbf{1}^T(\boldsymbol{l} - \boldsymbol{l^*}) = 0 \qquad \mathbf{1}^T(\boldsymbol{r} - \boldsymbol{r^*}) = 0$$

**Proposition 2.** *Residuals and estimated components are uncorrelated:*

$$\boldsymbol{m_0^{*T}}(\boldsymbol{m_1} - \boldsymbol{m_1^*}) = 0 \qquad \boldsymbol{m_0^{*T}}(\boldsymbol{m_2} - \boldsymbol{m_2^*}) = 0 \qquad \boldsymbol{m_0^{*T}}(\boldsymbol{l} - \boldsymbol{l^*}) = 0$$
$$\boldsymbol{m_0^{*T}}(\boldsymbol{r} - \boldsymbol{r^*}) = 0 \qquad \boldsymbol{m_1^{*T}}(\boldsymbol{m_2} - \boldsymbol{m_2^*}) = 0 \qquad \boldsymbol{m_1^{*T}}(\boldsymbol{l} - \boldsymbol{l^*}) = 0$$
$$\boldsymbol{m_1^{*T}}(\boldsymbol{r} - \boldsymbol{r^*}) = 0 \qquad \boldsymbol{m_2^{*T}}(\boldsymbol{l} - \boldsymbol{l^*}) = 0 \qquad \boldsymbol{m_2^{*T}}(\boldsymbol{r} - \boldsymbol{r^*}) = 0$$

**Proposition 3.** *Estimations and proper residuals are uncorrelated:*

$$\boldsymbol{m_0^{*T}}(\boldsymbol{m_0} - \boldsymbol{m_0^*}) = 0 \qquad \boldsymbol{m_1^{*T}}(\boldsymbol{m_1} - \boldsymbol{m_1^*}) = 0$$
$$\boldsymbol{m_2^{*T}}(\boldsymbol{m_2} - \boldsymbol{m_2^*}) = 0 \qquad \boldsymbol{l^{*T}}(\boldsymbol{l} - \boldsymbol{l^*}) = 0 \qquad \boldsymbol{r^{*T}}(\boldsymbol{r} - \boldsymbol{r^*}) = 0$$

**Proposition 4.** *Sums of residuals are zero:*

$$\mathbf{1}^T(\boldsymbol{\mu_{m_0}} - \boldsymbol{\mu^*_{m_0}}) = 0 \qquad \mathbf{1}^T(\boldsymbol{\mu_{m_1}} - \boldsymbol{\mu^*_{m_1}}) = 0 \qquad \mathbf{1}^T(\boldsymbol{\mu_{m_2}} - \boldsymbol{\mu^*_{m_2}}) = 0$$

**Proposition 5.** *Residuals and estimated components are uncorrelated:*

$$\boldsymbol{\mu^*_{m_0}}^T(\boldsymbol{\mu_{m_1}} - \boldsymbol{\mu^*_{m_1}}) = 0 \qquad \boldsymbol{\mu^*_{m_0}}^T(\boldsymbol{\mu_{m_2}} - \boldsymbol{\mu^*_{m_2}}) = 0$$
$$\boldsymbol{\mu^*_{m_1}}^T(\boldsymbol{\mu_{m_2}} - \boldsymbol{\mu^*_{m_2}}) = 0$$

**Proposition 6.** *Estimations and proper residuals are uncorrelated:*

$$\boldsymbol{\mu^*_{m_0}}^T(\boldsymbol{\mu_{m_0}} - \boldsymbol{\mu^*_{m_0}}) = 0 \qquad \boldsymbol{\mu^*_{m_1}}^T(\boldsymbol{\mu_{m_1}} - \boldsymbol{\mu^*_{m_1}}) = 0$$
$$\boldsymbol{\mu^*_{m_2}}^T(\boldsymbol{\mu_{m_2}} - \boldsymbol{\mu^*_{m_2}}) = 0$$

### 4.3.4 Goodness of fit indices for the model

In order to evaluate the goodness of fit for our model, we considered the standard $R^2$ and $R^2_{adj}$ (adjusted $R^2$). However, because we are working with the hypothesis of independence between $\tilde{\boldsymbol{Y}}^*_{\mathcal{M}_1}$ and $\tilde{\boldsymbol{Y}}^*_{\mathcal{M}_2}$ we separately considered the performances measured by the two fit indices. For the first model $\tilde{\boldsymbol{Y}}^*_{\mathcal{M}_1}$, the *total sum of squares* is:

$$\begin{aligned} \text{TSS} = &\|\boldsymbol{m_0} - \mathbf{1}\overline{\boldsymbol{m_0}}\|^2 + \|\boldsymbol{m_1} - \mathbf{1}\overline{\boldsymbol{m_1}}\|^2 + \\ &+ \|\boldsymbol{m_2} - \mathbf{1}\overline{\boldsymbol{m_2}}\|^2 + \|\boldsymbol{l} - \mathbf{1}\bar{\boldsymbol{l}}\|^2 + \|\boldsymbol{r} - \mathbf{1}\overline{\boldsymbol{r}}\|^2 \end{aligned} \tag{4.20}$$

whereas the *explained sum of squares* for the regression model is:

$$\begin{aligned}\text{ESS} = &\|\boldsymbol{m}_0^* - \mathbf{1}\overline{\boldsymbol{m}_0}\|^2 + \|\boldsymbol{m}_1^* - \mathbf{1}\overline{\boldsymbol{m}_1}\|^2 + \\ &+ \|\boldsymbol{m}_2^* - \mathbf{1}\overline{\boldsymbol{m}_2}\|^2 + \|\boldsymbol{l}^* - \mathbf{1}\bar{\boldsymbol{l}}\|^2 + \|\boldsymbol{r}^* - \mathbf{1}\bar{\boldsymbol{r}}\|^2\end{aligned} \tag{4.21}$$

finally, the *residuals sum of squares* is:

$$\begin{aligned}\text{RSS} = &\|\boldsymbol{m}_0 - \boldsymbol{m}_0^*\|^2 + \|\boldsymbol{m}_1 - \boldsymbol{m}_1^*\|^2 + \|\boldsymbol{m}_2 - \boldsymbol{m}_2^*\|^2 + \\ &+ \|\boldsymbol{l} - \boldsymbol{l}^*\|^2 + \|\boldsymbol{r} - \boldsymbol{r}^*\|^2\end{aligned} \tag{4.22}$$

Finally the following identity holds from the earlier definitions:

$$\text{TSS} = \text{ESS} + \text{RSS}$$

Its proof is trivial and therefore we omit it in this article. In other words, the above statement means that the total sum of squares is equal to the sum of the explained sum of squares and the residual sum of squares for each component of the model. In this context, the following index can be written:

$$R^2 = \frac{\text{ESS}}{\text{TSS}} \tag{4.23}$$

thus providing an interpretation which in line with the traditional regression approach for crisp data. The adjusted version for the $R^2$ is given as follows:

$$R^2_{adj} = 1 - (1 - R^2) \cdot \left(\frac{n-1}{n-q}\right) \quad \text{with: } q = (k+1) + 8 \tag{4.24}$$

where $q$ indicates the number of estimated parameters for the first component.

Similarly, for the second component $\tilde{\boldsymbol{Y}}^*_{\mathcal{M}_2}$ $R^2$ is written as in the previous case, where its components are the following:

$$\begin{aligned}\text{TSS} = &\|\boldsymbol{\mu}_{m_0} - \mathbf{1}\overline{\boldsymbol{\mu}}_{m_0}\|^2 + \|\boldsymbol{\mu}_{m_1} - \mathbf{1}\overline{\boldsymbol{\mu}}_{m_1}\|^2 + \\ &+ \|\boldsymbol{\mu}_{m_2} - \mathbf{1}\overline{\boldsymbol{\mu}}_{m_2}\|^2\end{aligned} \tag{4.25}$$

$$\begin{aligned}\text{ESS} = &\|\boldsymbol{\mu}^*_{m_0} - \mathbf{1}\overline{\boldsymbol{\mu}}_{m_0}\|^2 + \|\boldsymbol{\mu}^*_{m_1} - \mathbf{1}\overline{\boldsymbol{\mu}}_{m_1}\|^2 \\ &+ \|\boldsymbol{\mu}^*_{m_2} - \mathbf{1}\overline{\boldsymbol{\mu}}_{m_2}\|^2\end{aligned} \tag{4.26}$$

$$\begin{aligned} \text{RSS} = &\|\boldsymbol{\mu}_{m_0} - \boldsymbol{\mu}^*_{m_0}\|^2 + \|\boldsymbol{\mu}_{m_1} - \boldsymbol{\mu}^*_{m_1}\|^2 + \\ &+ \|\boldsymbol{\mu}_{m_2} - \boldsymbol{\mu}^*_{m_2}\|^2 \end{aligned} \tag{4.27}$$

the adjusted version of the $R^2$ index is:

$$R^2_{adj} = 1 - (1 - R^2) \cdot \left(\frac{n-1}{n-p}\right) \quad \text{with: } p = (k+1) + 4 \tag{4.28}$$

where $p$ is the number of estimated parameters for the second component. Note that, if we run an analysis on a small dataset with very few observations, then one should adjust the $R^2$ index only for a factor $q/p = n-(k+1)$.

### 4.3.5 Some remarks

**About the model complexity issue.**

It is important to note that the generative model requires $2(k+1)+12$ parameters ($[k+1]+8$ for $\tilde{\boldsymbol{Y}}^*_{\mathcal{M}_1}$ and $[k+1]+4$ for $\tilde{\boldsymbol{Y}}^*_{\mathcal{M}_2}$), whereas the corresponding non-generative model needs a total of $12 \cdot (k+1)$ parameters. Therefore, the generative-model is clearly more parsimonious than the non-generative one.

**About the model representation issue.**

We recall that our generative model is in line with a semi-confirmatory approach which assumes full consistency between data and the generative hypothesis. However, in some applicative contexts this hypothesis may not be sustainable and, therefore, an alternative approach based on *direct estimation* would instead be preferred (D'Urso, 2003). In particular, this non-generative representation might be suitable when generative hypothesis cannot be properly inferred on the basis of empirical data.

**About the algorithmic issue.**

In fitting the unconstrained fuzzy regression model, we adopted an iterative procedure based on standard stopping criteria and random initialization. However, one possible limitation of the unconstrained version of the

algorithm is that, in some circumstances, it might not yield feasible solutions. In particular, if the fuzzy regression model is fitted to empirical data which are largely corrupted by noise, the best parameters estimates might incur in violations of the *natural constraints* of the 2-mode fuzzy numbers (namely $\boldsymbol{m}_1^* < \boldsymbol{m}_0^* < \boldsymbol{m}_2^*$, $\boldsymbol{l}^* > \mathbf{0}_n$, $\boldsymbol{r}^* > \mathbf{0}_n$ and $\boldsymbol{\mu}_{m_1}^* < \boldsymbol{\mu}_{m_0}^* < \boldsymbol{\mu}_{m_2}^*$). In these situations, we would prefer a *constrained version* of the algorithm based on non-linear constrained programming optimization techniques (Gill, Murray, and Wright, 1981) or other optimization rationales like, for example, those at the base of the work of Lima Neto and De Carvalho (2010).

## 4.4 Applications

In this section we describe five applications to illustrate the main features of our fuzzy regression model. All the algorithms developed for these contributions are available upon request to the authors.

### 4.4.1 Example 1: Response Times and Moral Dilemma

In this first example we studied the relation between response time (crisp) and moral judgement (fuzzy). In cognitive decision making (Greene and Haidt, 2002; Haidt, 2001; Trevino, 1986), moral judgements and dilemmas are relevant phenomena characterized by high levels of uncertainty in individuals' responses. In this application, we used a moral dilemma based on the well-known *trolley scenario* (McGuire et al., 2009; Nichols and Mallon, 2006; Rai and Holyoak, 2010):

> *A trolley is running out of control down a track. In its path are five people who have been tied to the track by a madman. Fortunately, you can flip a switch that will lead the trolley down a different track to safety. Unfortunately, there is a single person tied to that track. Do you agree to flip the switch?*

In general, individuals react to the dilemma by using their personal moral beliefs. However, some individuals in solving the moral dilemma may hesitate in providing the final response. The trolley dilemma was administered to a group of students from the University of Trento (Italy)

and the responses were collected using a computerized interface based on the Mouse Tracking methodology (Freeman and Ambady, 2010; Johnson et al., 2012). In particular, participants were told that a pseudo-circular scale with five response levels (strongly disagree, disagree, neither, agree, strongly agree) would be presented on the screen, and that they were asked to choose which of these responses was the most appropriate for the presented dilemma. After participants clicked a start button, a window with the moral dilemma appeared at the top of the screen. Next the scale with the five levels appeared while the cursor was allocated to the center of the screen. Participants give their responses by mouse-clicking the chosen level of the scale. Meanwhile, we recorded the streaming x-y coordinates of the computer mouse (sampling rate of approximately 70 Hz). Figure 4.4 shows two empirical patterns of mouse movements with the associated contour density plots. In particular, figure 4.4-a represents an empirical pattern with a low uncertainty/fuzziness, by contrast figure 4.4-b shows a pattern with a higher level of uncertainty. Figure 4.5 shows the histograms and the associated fuzzy sets constructed using the radial positions of the x-y mouse movement coordinates of the empirical patterns. In particular, the fuzzy sets were obtained by an heuristic optimization procedure that allow to convert histograms into fuzzy sets (Medasani, Kim, and Krishnapuram, 1998; Nieradka and Butkiewicz, 2007). Table 4.1 shows the empirical data which were used for the regression analysis.

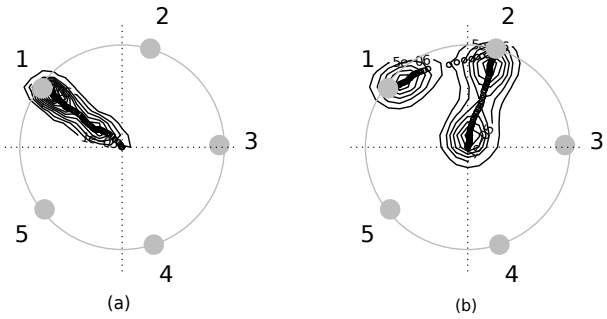



FIGURE 4.4: Mouse Tracker: empirical patterns of mouse movements (black circles) and their density (black contour lines). Note that the two patterns are different while sharing a same finale response (1 = strongly disagree). The numbers encode the five levels of the scale: 1 = strongly disagree, 2 = disagree, 3 = neither, 4 = agree, 5 = strongly agree).

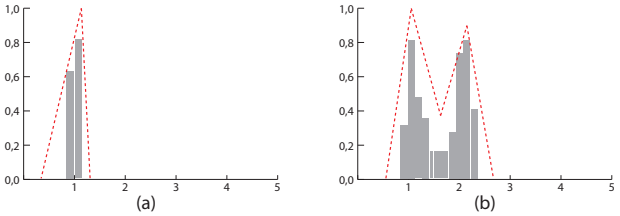



FIGURE 4.5: Mouse Tracker: histograms of the mouse movements represented in figure 4.4 and the related fuzzy sets (in dotted red line). Note that the histograms were rescaled in order to provide a better comparison with the fuzzy sets.

TABLE 4.1: Example 1: Original Dataset

| $x$ | $m_1$ | $m_0$ | $m_2$ | $l$ | $r$ | $\mu_{m_1}$ | $\mu_{m_0}$ | $\mu_{m_2}$ |
|---|---|---|---|---|---|---|---|---|
| 1.18 | 0.70 | 0.97 | 1.13 | 0.25 | 0.22 | 1.00 | 0.80 | 1.00 |
| 2.11 | 0.70 | 0.95 | 1.13 | 0.27 | 0.20 | 1.00 | 0.80 | 1.00 |
| 3.10 | 0.73 | 0.98 | 1.18 | 0.33 | 0.17 | 1.00 | 0.75 | 1.00 |
| 4.10 | 0.84 | 1.73 | 2.38 | 0.34 | 0.27 | 1.00 | 0.65 | 1.00 |
| 5.27 | 0.82 | 1.73 | 2.13 | 0.24 | 0.27 | 1.00 | 0.65 | 0.88 |
| 6.32 | 0.80 | 1.65 | 2.64 | 0.28 | 0.21 | 1.00 | 0.70 | 0.84 |
| 7.51 | 0.83 | 1.73 | 2.38 | 0.25 | 0.17 | 0.70 | 0.60 | 1.00 |
| 8.45 | 0.70 | 1.85 | 2.25 | 0.25 | 0.28 | 0.85 | 0.80 | 1.00 |
| 9.42 | 0.83 | 2.10 | 3.64 | 0.33 | 0.27 | 1.00 | 0.70 | 1.00 |
| 11.03 | 0.80 | 2.23 | 3.64 | 0.25 | 0.16 | 0.78 | 0.60 | 1.00 |
| 12.01 | 4.59 | 5.37 | 6.02 | 0.32 | 0.28 | 0.77 | 0.40 | 1.00 |

For the first component of the model, the regression algorithm converged after 113 iterations, whereas for the second component it required 132 iterations. The estimated values and parameters are shown in tables 4.2 and 4.3. As table 4.2 and figure 4.6 show, all the estimated components satisfy the conditions (c.4-c.10) characterizing the 2-mode fuzzy numbers. The performance of the model was good ($R^2 = 0.68$ for the first component; $R^2 = 0.45$ for the second component).

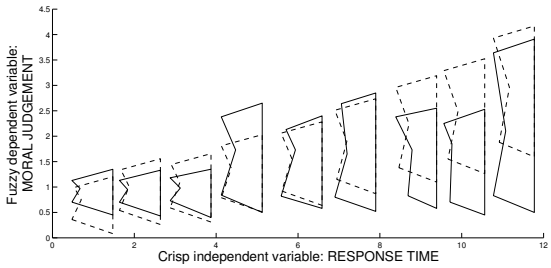



FIGURE 4.6: Example 1: Scatterplot with empirical (solid line) and estimated fuzzy numbers (dashed line)

The results show a slight increasing relation between response times and

TABLE 4.2: Example 1: Estimated values

| $m_1^*$ | $m_0^*$ | $m_2^*$ | $l^*$ | $r^*$ | $\mu^*_{m_1}$ | $\mu^*_{m_0}$ | $\mu^*_{m_2}$ |
|---|---|---|---|---|---|---|---|
| 0.36 | 0.75 | 0.99 | 0.25 | 0.22 | 1.00 | 0.77 | 0.94 |
| 0.54 | 1.01 | 1.34 | 0.27 | 0.20 | 1.00 | 0.76 | 0.97 |
| 0.59 | 1.09 | 1.44 | 0.33 | 0.17 | 1.00 | 0.76 | 0.97 |
| 0.79 | 1.37 | 1.81 | 0.34 | 0.27 | 1.00 | 0.76 | 1.00 |
| 0.92 | 1.56 | 2.06 | 0.24 | 0.27 | 0.98 | 0.74 | 1.00 |
| 1.15 | 1.90 | 2.51 | 0.28 | 0.21 | 0.95 | 0.71 | 1.00 |
| 1.38 | 2.24 | 2.96 | 0.25 | 0.17 | 0.92 | 0.68 | 1.00 |
| 1.55 | 2.49 | 3.29 | 0.25 | 0.28 | 0.90 | 0.65 | 1.00 |
| 1.88 | 2.96 | 3.93 | 0.33 | 0.27 | 0.86 | 0.61 | 1.00 |
| 2.03 | 3.17 | 4.20 | 0.25 | 0.16 | 0.84 | 0.59 | 1.00 |
| 2.43 | 3.76 | 4.98 | 0.32 | 0.28 | 0.79 | 0.53 | 1.00 |

TABLE 4.3: Example 1: Parameters for the model

| | | | | | |
|---|---|---|---|---|---|
| $\hat{\boldsymbol{a}}$ = | 0.25 | -0.26 | | $\hat{\boldsymbol{\alpha}}$ = | -0.34 0.84 |
| $\hat{b}$ = | 1.00 | $\hat{r_1}$ = | -0.07 | $\hat{\beta}$ = | -0.10 |
| $\hat{c}$ = | 1.76 | $\hat{r_2}$ = | 0.50 | $\hat{\gamma}$ = | 1.04 |
| $\hat{e}$ = | 1.35 | $\hat{r_3}$ = | 0.27 | $\xi_1$ = | 0.27 |
| $\hat{d}$ = | 1.31 | $\hat{r_4}$ = | 0.08 | $\xi_2$ = | 1.03 |

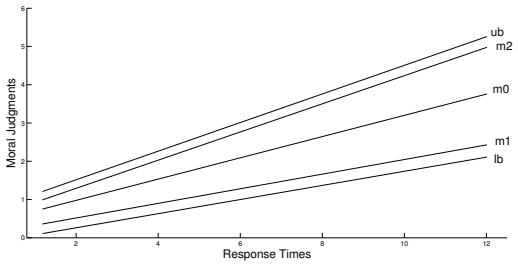



FIGURE 4.7: Example 1: Fuzzy regression plot

final responses ($\hat{\boldsymbol{a}} = 0.25$). By contrast, we observed a strong increasing relation between the final responses and fuzziness ($\hat{c} = 1.76$ and $\hat{d} = 1.31$). In particular, the slower the response time, the greater the uncertainty/fuzziness of the final response.

### 4.4.2 Example 2: Private Consumption and Perceived Satisfaction

In this second example we studied the relation between private consumption (crisp) and perceived life satisfaction (fuzzy). In this application, data were collected by means of a computerized questionnaire based on fuzzy rating scales. This procedure is widely used in human ratings or

human evaluations studies (Gil and González-Rodríguez, 2012; Hesketh et al., 1989; Hesketh, Pryor, and Hesketh, 1988; Lalla, Facchinetti, and Mastroleo, 2005). In particular, the fuzzy scale is based on a pseudo-continuous scale representing qualitative ordinal values (see figure 4.8).[4] By using a graphical interface, the respondent moves the mouse cursor toward the preferred level. In this context, two general scenarios can arise: the respondent chooses a single level of the scale (Figure 4.9-a) or s(he) selects an intermediate position which lies between the two levels (Figure 4.9-b).

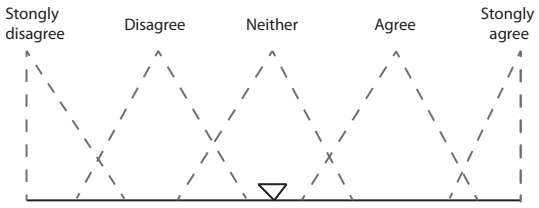



FIGURE 4.8: Fuzzy rating scale with a pseudo-continuous line and a movable cursor (in solid line) and a hidden fuzzy variable (in dotted line) representing the levels of the scale.

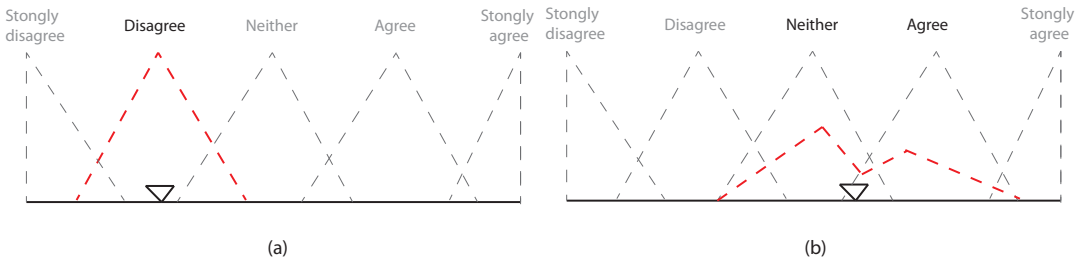



FIGURE 4.9: Fuzzy rating scale: (a) convex-fuzzy set and (b) non-convex fuzzy set, associated to the respondent's responses. In grey dotted lines are represented the levels of the scale while in red dotted lines the activated final levels.

The scale was administered to a group of students from the University of Trento (Italy). For descriptive purposes, in this application we selected only a small group of units characterized by non-convex features (see Table 5.1).

For the first component of the model, the algorithm converged after 138

[4]The fuzzy sets were obtained from the histograms of the empirical responses by maximizing the entropy of the data (Avci and Avci, 2009; Cheng and Chen, 1997; Medasani, Kim, and Krishnapuram, 1998; Nieradka and Butkiewicz, 2007).

TABLE 4.4: Example 2: Original Dataset

| $x$ | $m_1$ | $m_0$ | $m_2$ | $l$ | $r$ | $\mu_{m_1}$ | $\mu_{m_0}$ | $\mu_{m_2}$ |
|---|---|---|---|---|---|---|---|---|
| 1.00 | 10.00 | 16.00 | 18.00 | 2.00 | 1.00 | 1.00 | 0.05 | 0.60 |
| 2.00 | 10.00 | 13.00 | 15.00 | 2.00 | 1.00 | 1.00 | 0.11 | 0.20 |
| 3.00 | 7.00 | 9.00 | 13.00 | 1.00 | 1.00 | 0.20 | 0.16 | 1.00 |
| 4.00 | 5.00 | 8.00 | 10.00 | 2.00 | 1.00 | 0.40 | 0.19 | 1.00 |
| 5.00 | 5.00 | 9.00 | 11.00 | 2.00 | 1.00 | 0.40 | 0.16 | 1.00 |
| 6.00 | 6.00 | 7.00 | 9.00 | 2.00 | 1.00 | 0.40 | 0.18 | 1.00 |
| 7.00 | 1.00 | 5.00 | 10.00 | 1.00 | 1.00 | 0.31 | 0.26 | 1.00 |
| 8.00 | 3.00 | 4.00 | 8.00 | 1.00 | 1.00 | 0.52 | 0.47 | 1.00 |
| 9.00 | 2.00 | 3.00 | 5.00 | 1.00 | 4.00 | 1.00 | 0.50 | 1.00 |
| 10.00 | 1.00 | 3.00 | 5.00 | 1.00 | 4.00 | 0.75 | 0.69 | 1.00 |

iterations while for the second component of the model, the algorithm required some additional iterations (202) to converge. The estimated values and parameters are shown in tables 4.5 and 4.6, respectively.

TABLE 4.5: Example 2: Estimated values

| $m_1^*$ | $m_0^*$ | $m_2^*$ | $l^*$ | $r^*$ | $\mu_{m_1}^*$ | $\mu_{m_0}^*$ | $\mu_{m_2}^*$ |
|---|---|---|---|---|---|---|---|
| 9.58 | 13.78 | 16.24 | 2.02 | 0.29 | 0.99 | 0.02 | 1.00 |
| 8.56 | 12.43 | 14.94 | 1.90 | 0.58 | 0.90 | 0.09 | 1.00 |
| 7.55 | 11.08 | 13.64 | 1.79 | 0.87 | 0.83 | 0.16 | 1.00 |
| 6.53 | 9.73 | 12.35 | 1.67 | 1.16 | 0.76 | 0.23 | 1.00 |
| 5.51 | 8.38 | 11.05 | 1.56 | 1.45 | 0.71 | 0.29 | 1.00 |
| 4.49 | 7.02 | 9.75 | 1.44 | 1.75 | 0.66 | 0.34 | 1.00 |
| 3.47 | 5.67 | 8.45 | 1.33 | 2.04 | 0.61 | 0.38 | 1.00 |
| 2.45 | 4.32 | 7.16 | 1.21 | 2.33 | 0.57 | 0.42 | 1.00 |
| 1.44 | 2.97 | 5.86 | 1.10 | 2.62 | 0.54 | 0.46 | 1.00 |
| 0.42 | 1.62 | 4.56 | 0.98 | 2.91 | 0.51 | 0.49 | 1.00 |

TABLE 4.6: Example 2: Parameters for the model

| | | | | | |
|---|---|---|---|---|---|
| $\hat{\boldsymbol{a}}$ = | -1.35 | 15.13 | $\hat{\boldsymbol{\alpha}}$ = | 0.62 | -0.06 |
| $\hat{b}$ = | 0.75 | $\hat{r_1}$ = -0.80 | $\hat{\beta}$ = | -0.09 | |
| $\hat{c}$ = | 0.11 | $\hat{r_2}$ = 0.93 | $\hat{\gamma}$ = | 0.90 | |
| $\hat{e}$ = | 0.96 | $\hat{r_3}$ = 3.01 | $\xi_1$ = | 0.62 | |
| $\hat{d}$ = | -0.22 | $\hat{r_4}$ = 3.93 | $\xi_2$ = | 0.63 | |

Notice that, the estimated components described in table 4.5 and represented in figure 4.10 satisfy the conditions (c.4-c.10) of the 2-mode fuzzy numbers. The performances of the model was very good for the first component of the model ($R^2 = 0.88$) and discrete for the second one

($R^2 = 0.40$). Figures 4.10 and 4.11 show the results of the fuzzy regression model.

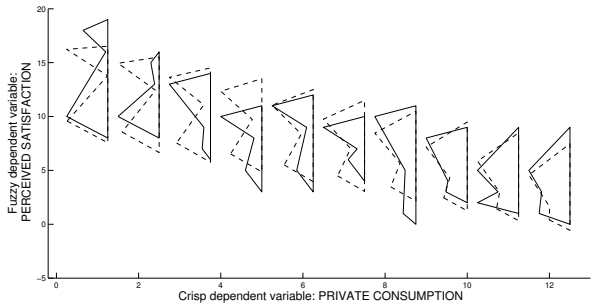



FIGURE 4.10: Example 2: Scatterplot with empirical (solid line) and estimated fuzzy numbers (dashed line)

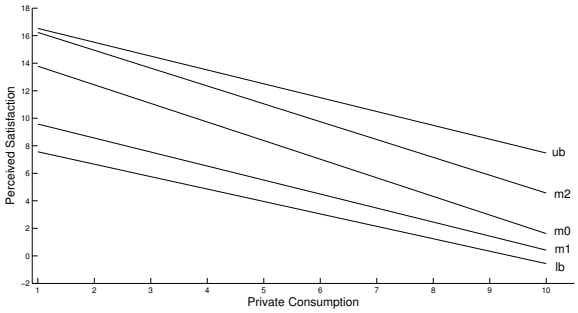



FIGURE 4.11: Example 2: Fuzzy regression plot

We observed a negative relation between perceived satisfaction and private consumption ($\hat{\boldsymbol{a}} = -1.35$). Moreover, there was a generative relation between $\boldsymbol{m_0^*}$ and $\boldsymbol{m_1^*}$ ($\hat{b} = 0.75$) as well as between $\boldsymbol{m_0^*}$ and $\boldsymbol{m_2^*}$ ($\hat{e} = 0.96$). Left spread had a weak decreasing relation with $\boldsymbol{m_1^*}$ ($\hat{c} = 0.11$) whereas right spread showed an increasing relation with $\boldsymbol{m_2^*}$ ($\hat{d} = -0.22$). Finally, the total spread of the perceived satisfaction increased as a function of increasing levels of private consumption. In summary, perceived satisfaction seemed to be inversely related to private consumptions.

### 4.4.3 Example 3: Unemployment and Employment rates

In this example we tested our model on a dataset with two crisp independent variables (Unemployment rate for 2010 and Employment rate for 2011) and one fuzzy dependent variable (Unemployment rate for 2011). Data were collected by OECD (OECD, 2011; OECD, 2013) and referred to a

subset of European countries (see Table 4.7).[5] In order to test our model in the case of multivariate analysis, the three crisp variables were re-scaled to a common scale and the dependent variable was also fuzzified with a suitable fuzzification procedure based on traditional Mamdani fuzzy system. The fuzzification routine yielded the following sets for the fuzzy dependent variable: *very-low* (1,2,5), *low* (2,6,10), *middle* (6,11,12), *high* (11,12,15), *very-high* (12,15,16).

TABLE 4.7: Example 3: Original Dataset

| | $x_1$ | $x_2$ | $m_1$ | $m_0$ | $m_2$ | $l$ | $r$ | $\mu_{m_1}$ | $\mu_{m_0}$ | $\mu_{m_2}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| B | 5.30 | 5.82 | 2.00 | 3.00 | 6.00 | 1.00 | 3.00 | 0.67 | 0.44 | 1.00 |
| C | 4.36 | 8.71 | 2.00 | 3.00 | 6.00 | 1.00 | 3.00 | 0.67 | 0.44 | 1.00 |
| E | 13.14 | 8.28 | 6.00 | 7.00 | 11.00 | 4.00 | 0.00 | 0.42 | 0.31 | 1.00 |
| D | 4.18 | 13.90 | 2.00 | 5.00 | 6.00 | 1.00 | 3.00 | 1.00 | 0.28 | 0.37 |
| G | 9.17 | 1.00 | 12.00 | 14.00 | 15.00 | 1.00 | 0.00 | 1.00 | 0.33 | 0.50 |
| H | 7.90 | 1.20 | 6.00 | 10.00 | 11.00 | 4.00 | 0.00 | 1.00 | 0.21 | 0.27 |
| P | 7.72 | 7.57 | 6.00 | 7.00 | 11.00 | 4.00 | 0.00 | 0.42 | 0.31 | 1.00 |
| K | 10.92 | 3.98 | 6.00 | 7.00 | 11.00 | 4.00 | 0.00 | 0.42 | 0.31 | 1.00 |
| S | 5.36 | 15.07 | 2.00 | 3.00 | 6.00 | 1.00 | 3.00 | 0.67 | 0.44 | 1.00 |

For the first component of the model, the algorithm converged after 331 iterations whereas for the second component the algorithm converged after 242 iterations. The estimated values and parameters are shown in tables 4.8 and 5.2, respectively.

TABLE 4.8: Example 3: Estimated values

| | $m_1^*$ | $m_0^*$ | $m_2^*$ | $l^*$ | $r^*$ | $\mu_{m_1}^*$ | $\mu_{m_0}^*$ | $\mu_{m_2}^*$ |
|---|---|---|---|---|---|---|---|---|
| B | 4.36 | 6.01 | 8.66 | 2.14 | 1.61 | 1.00 | 0.47 | 0.81 |
| C | 2.95 | 4.55 | 7.14 | 1.61 | 2.35 | 1.00 | 0.48 | 0.86 |
| E | 7.08 | 8.83 | 11.58 | 3.15 | 0.19 | 0.39 | 0.23 | 1.00 |
| D | 1.09 | 2.63 | 5.15 | 0.92 | 3.32 | 0.87 | 0.43 | 1.00 |
| G | 7.77 | 9.53 | 12.31 | 3.40 | 0.00 | 1.00 | 0.48 | 0.94 |
| H | 7.12 | 8.86 | 11.62 | 3.16 | 0.17 | 1.00 | 0.48 | 0.83 |
| P | 4.86 | 6.53 | 9.19 | 2.32 | 1.35 | 0.84 | 0.42 | 1.00 |
| K | 7.54 | 9.30 | 12.07 | 3.32 | 0.0 | 0.72 | 0.37 | 1.00 |
| S | 1.23 | 2.77 | 5.29 | 0.97 | 3.25 | 0.70 | 0.36 | 1.00 |

[5]We used the following abbreviations, B = Belgium, C = Czech Republic, E = Estonia, D = Germany, G = Greece, H = Hungary, P = Portugal, K = Slovak Republic, S = Sweden.

TABLE 4.9: Example 3: Parameters for the model

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\hat{\boldsymbol{a}}$ = | 0.47 | -0.35 | 5.57 | $\hat{\boldsymbol{\alpha}}$ = | 0.23 | -0.14 | 0.54 |
| $\hat{b}$ = | 0.97 | $\hat{r_1}$ = | -1.48 | $\hat{\beta}$ = | 2.72 | | |
| $\hat{c}$ = | 0.37 | $\hat{r_2}$ = | 0.52 | $\hat{\gamma}$ = | -2.23 | | |
| $\hat{e}$ = | 1.04 | $\hat{r_3}$ = | 2.42 | $\xi_1$ = | -0.24 | | |
| $\hat{d}$ = | -0.49 | $\hat{r_4}$ = | 5.83 | $\xi_2$ = | 1.50 | | |

As in the previous example, the estimated components described in table 4.8 satisfy the conditions (c.4-c.10) underlying the 2-mode fuzzy numbers. The performance of the model was good for the first component ($R^2 = 0.64$) and modest for the second component ($R^2 = 0.21$). As expected, the unemployment rate of the previous year (2010) was a good predictor for the unemployment rate of the subsequent year (2011), with $\hat{a}_{x_1} = 0.47$. Moreover, $\boldsymbol{m_1^*}$ and $\boldsymbol{m_2^*}$ had a good generative relation with $\boldsymbol{m_0^*}$ ($\hat{b} = 0.97$ and $\hat{e} = 1.04$). A similar pattern was observed for the left and right spreads ($\hat{c} = 0.37$ and $\hat{d} = -0.49$). In sum, unemployment seemed to increase as a function of decreasing levels of spread, that is to say, there was a very little uncertainty in classifying some countries into the high unemployment country cluster.

### 4.4.4 Example 4: Degenerate non-convex fuzzy data - triangular case

In this example we tested our model with degenerate non-convex fuzzy data based on standard triangular shapes. The data referred to the example studied by (Bisserier, Boukezzoula, and Galichet, 2010) and are described in Table 4.10. Note that in this example the second model is a trivial one and, therefore, we will not discuss it here.

The performance of the algorithm was excellent and required 144 iterations to converge. Estimated data and parameters are shown in tables 4.11 and 4.12, respectively.

Figures 4.12 and 4.13 show the graphical results of this application.

As expected, the results showed that our model correctly estimated triangular fuzzy numbers. In particular, note as $\hat{b} = \hat{e} = 1$, $\hat{r_1} = \hat{r_3} = 0$, $\boldsymbol{h}^* = \boldsymbol{i}^* = \mathbf{0}_n$ and how left $\boldsymbol{l}^*$ and right $\boldsymbol{r}^*$ spreads directly refer to $\boldsymbol{m_0^*}$.

TABLE 4.10: Example 4: Original Dataset

| $x$ | $m_1$ | $m_0$ | $m_2$ | $l$ | $r$ |
|---|---|---|---|---|---|
| 0.10 | 2.25 | 2.25 | 2.25 | 0.75 | 0.75 |
| 0.20 | 2.88 | 2.88 | 2.88 | 0.88 | 0.88 |
| 0.30 | 2.50 | 2.50 | 2.50 | 1.00 | 1.00 |
| 0.40 | 4.25 | 4.25 | 4.25 | 1.75 | 1.75 |
| 0.50 | 4.00 | 4.00 | 4.00 | 1.50 | 1.50 |
| 0.60 | 5.25 | 5.25 | 5.25 | 1.25 | 1.25 |
| 0.70 | 7.50 | 7.50 | 7.50 | 2.00 | 2.00 |
| 0.80 | 8.50 | 8.50 | 8.50 | 1.50 | 1.50 |

TABLE 4.11: Example 4: Estimated values

| $m_1^*$ | $m_0^*$ | $m_2^*$ | $l^*$ | $r^*$ |
|---|---|---|---|---|
| 1.52 | 1.52 | 1.52 | 0.86 | 0.86 |
| 2.41 | 2.41 | 2.41 | 0.99 | 0.99 |
| 3.30 | 3.30 | 3.30 | 1.13 | 1.13 |
| 4.20 | 4.20 | 4.20 | 1.26 | 1.26 |
| 5.09 | 5.09 | 5.09 | 1.40 | 1.40 |
| 5.98 | 5.98 | 5.98 | 1.53 | 1.53 |
| 6.87 | 6.87 | 6.87 | 1.67 | 1.67 |
| 7.76 | 7.76 | 7.76 | 1.80 | 1.80 |

TABLE 4.12: Example 4: Parameters for the model

| | | | |
|---|---|---|---|
| $\hat{\boldsymbol{a}} =$ | 0.89 | 0.63 | |
| $\hat{b} =$ | 1.00 | $\hat{r_1} =$ | 0.00 |
| $\hat{c} =$ | 0.15 | $\hat{r_2} =$ | 0.62 |
| $\hat{e} =$ | 1.00 | $\hat{r_3} =$ | 0.00 |
| $\hat{d} =$ | 0.15 | $\hat{r_4} =$ | 0.62 |

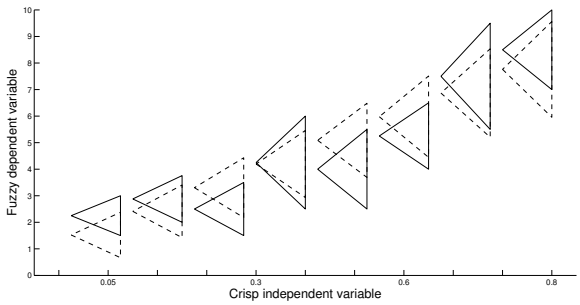



FIGURE 4.12: Example 4: Scatterplot with empirical (solid line) and estimated fuzzy numbers (dashed line)

### 4.4.5 Example 5: Degenerate non-convex fuzzy data - trapezoidal case

In this last application we tested our model on convex trapezoidal data formerly published in D'Urso (2003). For descriptive purposes, in this

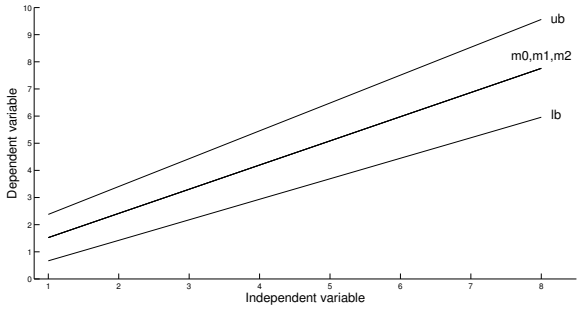



FIGURE 4.13: Example 4: Fuzzy regression plot

example we selected only two variables: the cost/km ratio index (as a crisp independent variable) and expert decision rating (as fuzzy dependent variable). Moreover, in order to show the performance of the fuzzy regression model, we selected 12 cases with clear trapezoidal shapes from the original dataset (Table 4.13).

TABLE 4.13: Example 5: Original Dataset

| $x$ | $m_1$ | $m_0$ | $m_2$ | $l$ | $r$ |
|---|---|---|---|---|---|
| 0.30 | 7.00 | 7.50 | 8.00 | 1.00 | 1.00 |
| 0.36 | 4.00 | 4.50 | 5.00 | 1.00 | 1.00 |
| 0.37 | 4.00 | 4.50 | 5.00 | 1.00 | 1.00 |
| 0.41 | 6.00 | 6.50 | 7.00 | 1.00 | 1.00 |
| 0.49 | 6.00 | 6.50 | 7.00 | 1.00 | 1.00 |
| 0.50 | 4.00 | 4.50 | 5.00 | 1.00 | 1.00 |
| 0.52 | 6.00 | 6.50 | 7.00 | 1.00 | 1.00 |
| 0.62 | 7.00 | 7.50 | 8.00 | 1.00 | 1.00 |
| 0.65 | 7.00 | 7.50 | 8.00 | 1.00 | 1.00 |
| 0.84 | 7.00 | 7.50 | 8.00 | 1.00 | 1.00 |
| 0.88 | 7.00 | 7.50 | 8.00 | 1.00 | 1.00 |
| 1.01 | 8.00 | 8.50 | 10.00 | 1.00 | 0.00 |

The resulting dataset and parameters are described in Tables 4.14 and 4.15. Figures 4.14 and 4.15 show the graphical results.

## 4.5 Conclusion and further perspectives

In this chapter we discussed the concept of non-convex fuzzy numbers as a possible extension of convex-fuzzy numbers. In what follows we briefly list some relevant aspects of our novel proposal.

TABLE 4.14: Example 5: Estimated values

| $m_1^*$ | $m_0^*$ | $m_2^*$ | $l^*$ | $r^*$ |
|---|---|---|---|---|
| 4.97 | 5.39 | 5.84 | 1.00 | 1.13 |
| 5.21 | 5.66 | 6.12 | 1.00 | 1.08 |
| 5.25 | 5.70 | 6.17 | 1.00 | 1.08 |
| 5.41 | 5.88 | 6.36 | 1.00 | 1.04 |
| 5.73 | 6.23 | 6.74 | 1.00 | 0.98 |
| 5.77 | 6.28 | 6.79 | 1.00 | 0.98 |
| 5.85 | 6.36 | 6.89 | 1.00 | 0.96 |
| 6.25 | 6.81 | 7.36 | 1.00 | 0.89 |
| 6.37 | 6.94 | 7.50 | 1.00 | 0.86 |
| 7.13 | 7.77 | 8.41 | 1.00 | 0.72 |
| 7.29 | 7.95 | 8.60 | 1.00 | 0.69 |
| 7.80 | 8.52 | 9.21 | 1.00 | 0.59 |

TABLE 4.15: Example 5: Parameters for the model

| | | | |
|---|---|---|---|
| $\hat{\boldsymbol{a}}$ = | 4.40 | 4.07 | |
| $\hat{b}$ = | 0.90 | $\hat{r_1}$ = | 0.08 |
| $\hat{c}$ = | 0.00 | $\hat{r_2}$ = | 1.00 |
| $\hat{e}$ = | 1.08 | $\hat{r_3}$ = | 0.02 |
| $\hat{d}$ = | -0.16 | $\hat{r_4}$ = | 2.06 |

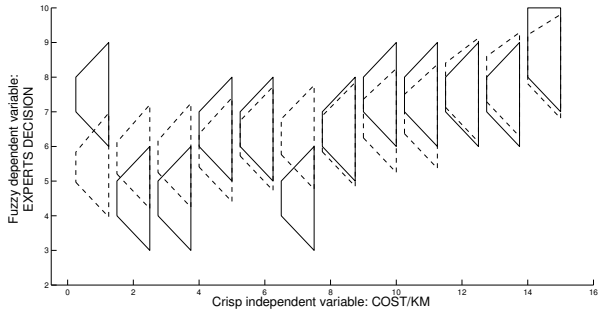



FIGURE 4.14: Example 5: Scatterplot with empirical (solid line) and estimated fuzzy numbers (dashed line)

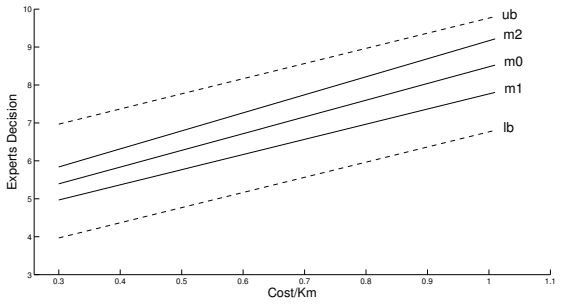



FIGURE 4.15: Example 5: Fuzzy regression plot

Among the positive aspects, we stress how the fuzzy regression model described in this contribution allows to i) generalize some known fuzzy regression models based on standard convex fuzzy data, ii) provide a good performance with non-convex data as well as convex (triangular as well as trapezoidal) data.
However, our model can potentially suffer from some limitations. For example, in some empirical phenomena the generative representation assumption and the corresponding independence assumption between the components cannot be valid. In particular, the unconstrained algorithm could estimate fuzzy numbers which violate their normative representation.
Various possible extensions of our proposal to represent non-convex fuzzy numbers could be considered, both from the algorithmic and the modeling perspectives. For example, we could think of a new a fuzzy regression model for non-convex data based on a two-step hierarchical estimation procedure in which the components are estimated in a nested fashion using a full global model. Moreover, the adoption of a constrained approach for the estimation procedure (e.g., non-linear constrained optimization or other approaches, such as Generalized Maximum Entropy methods of estimation (Ciavolino and Dahlgaard, 2009; Golan and Judge, 1996)) could always guarantee the consistency between the parameters estimated and the corresponding normative representations for non-convex fuzzy numbers. A future venue of research, would also consist in the generalization of the proposed approach to Fuzzy Random Variables (FRV) (Colubi et al., 2001; Gil, López-Díaz, and Ralescu, 2006; Viertl, 2011) for non-convex data. This would extend our proposal beyond the descriptive approach presented in this contribution.
Finally, it should not be difficult to extend our approach to consider i) the development of fuzzy-input fuzzy-output and fuzzy-input crisp-output regression models for non-convex data and ii) the extension of the regression model and its estimation procedure to cases where data are described by means of 2-mode fuzzy numbers beyond the ones characterized by

piecewise linear membership functions (e.g., gaussian 2-mode fuzzy numbers).

In sum, non-convexity seems to be an important feature in many empirical phenomena and in order to manage this kind of information, the development of suitable statistical approaches could be an important step forward for sciences devoted to manage with fuzziness and uncertainty.

# Chapter 5

# A dimension reduction technique for non-convex fuzzy data

The content of the chapter has to appear as: Calcagnì, A., Lombardi, L., & Pascali, E. (2014). A dimension reduction technique for two-mode non-convex fuzzy data. *Soft Computing*, 1-14, In Press

## 5.1 Introduction

In many research fields such as, for example, behavioural and social sciences, epidemiology, bioinformatics, engineering, researchers often have to deal with high dimensional datasets which are usually represented by $n$ (units) $\times$ $m$ (variables) matrices. Country statistical profiles, socio economic tables, chemical databases, survival tables, and self-report questionnaires, are all examples of this type of data structures (Eriksson, 2006). In such contexts, it may be useful to reduce the dimensionality, complexity, of these large datasets. This may happen, for instance, when a researcher wants to enhance the efficiency and accuracy of a data analysis, or when s(he) wants to extract the most relevant information from the available data. In all these cases, several *dimension reduction techniques* such as, for instance, Principal Component Analysis, Independent Component Analysis, Multidimensional Scaling, Cluster Analysis, and Latent

Semantic Analysis, are available to perform data analysis on large structures (Hastie, Tibshirani, and Friedman, 2001). Among these options, Principal Component Analysis (PCA) is a well-known and widely used unsupervised variables transformation technique for linear dimensionality reduction. Its aim is to summarize a $n \times m$ data matrix into a new $n \times p$ reduced model matrix (with $p \ll m$) which reconstructs the information contained in the original data (Abdi and Williams, 2010). Usually, PCA can be performed using different mathematical procedures such as, for example, eigenvalues decomposition, singular values decomposition, low-rank approximations, and component analysis.

PCA has been mainly applied to standard crisp data. However, some researchers have extended the PCA framework also to more complex data (e.g., interval, symbolic or fuzzy) to better model variables with vague and imprecise information (Lauro and Palumbo, 2000; Douzal-Chouakria, Billard, and Diday, 2011; Taheri, 2003; Viertl, 2011). A natural way to model imprecision and vagueness in empirical data is by means of the so-called fuzzy sets (Zimmermann, 2001).

Conventionally, fuzzy sets have been described by *LR-type representation* (Dubois et al., 1988) which is primarily used for modelling convex-shaped fuzzy objects. However, in some empirical contexts such as, for example, human decision making and ratings, convex representations might not be capable to capture more complex structures in the data. Moreover, non-convexity seems to arise as a natural property in many applications based on fuzzy systems, such fuzzy decision making and expert systems (Calcagnì, Lombardi, and Pascali, 2013; Garibaldi et al., 2004; Facchinetti and Pacchiarotti, 2006; Reuter, 2008). In this framework, the use of standard LR type representation could be questionable. A possible way out consists in adopting ad-hoc data manipulation procedures to transform non-convex data into standard convex representation (e.g., by using Graham Scan algorithm or Steiner symmetrization). However, one serious

limitation of these data transformation procedures is that they can artificially mask relevant information carried out by the non-convexity property. Unfortunately, reduction dimension techniques for analysing non-convex fuzzy data, as far as we know, has not been proposed yet in the literature. In this chapter we present a novel dimension reduction technique, called *non-convex fuzzy component analysis* (NCFCA), which is based on the frameworks of Component Analysis - CA (Meredith and Millsap, 1985; Millsap and Meredith, 1988) and least squares approach (Diamond, 1988; Giordani and Kiers, 2004a). Unlike other fuzzy modelling procedure, NCFCA always guarantees a direct modelling of multidimensional fuzzy data with possibly non-convex shapes.
The reminder of the chapter is organized as follows. The second section is devoted to briefly recall the basic characteristics of convex as well as non-convex fuzzy data. The third section exposes the component analysis for non-convex fuzzy data together with its main features. Moreover, this section also describes some useful procedures for data fitting and model evaluation. The fourth section illustrates three applications of the proposed method to some behavioural and socioeconomic data collected using different procedures (e.g., fuzzy scales of measurement and fuzzy measurement systems). Finally, the fifth section concludes this chapter providing some final remarks and suggestions for future extensions of our approach.

## 5.2 Non-convex fuzzy component analysis (NCFCA)

In this section we provide a detailed description of the NCFCA model. From a least squares viewpoint, the main idea is to reduce the dimensionality of the underlying structure of the non-convex fuzzy data by finding a set of *components* which minimize a specific distance between the empirical data and the fuzzy model data. For the sake of simplicity, in this article we describe a technique which is restricted to deal with piecewise

linear 2-mode representations and/or degenerated triangular and trapezoidal fuzzy data. Although some empirical contexts may require different representations for non-convex fuzzy data (e.g., quadratic 2-mode fuzzy numbers), in this contribution we introduce a dimension reduction technique for the most simple case first.

### 5.2.1 Model and data analysis

Let $\mathbf{X}$ be a $n$ (units) $\times$ $m$ (variables) data matrix representing the observed data. The generic element $x_{ij}$ of $\mathbf{X}$ defines the array

$$x_{ij} = \{m_0, h, i, l, r, \mu_0, \mu_1, \mu_2\}_{ij}$$

representing a parametrized fuzzy set. By adopting the parametric representation for 2-mode fuzzy data, the elements of $\mathbf{X}$ can be represented by a collection of $n \times m$ matrices, $\mathbf{M_0}$, $\mathbf{H}$, $\mathbf{I}$, $\mathbf{L}$, $\mathbf{R}$, $\mathbf{MU_0}$, $\mathbf{MU_1}$, $\mathbf{MU_2}$ which contain the set of parameters involved in the LHIR-representation. Therefore, the component model for 2-mode fuzzy data can be expressed as follows:

$$\begin{cases} \mathbf{M_0} = \mathbf{\Psi}_{M_0}\mathbf{\Gamma}^T + \mathbf{E}_{M_0} \\ \mathbf{H} = \mathbf{\Psi}_{H}\mathbf{\Gamma}^T + \mathbf{E}_{H} \\ \mathbf{I} = \mathbf{\Psi}_{I}\mathbf{\Gamma}^T + \mathbf{E}_{I} \\ \mathbf{L} = \mathbf{\Psi}_{L}\mathbf{\Gamma}^T + \mathbf{E}_{L} \\ \mathbf{R} = \mathbf{\Psi}_{R}\mathbf{\Gamma}^T + \mathbf{E}_{R} \\ \mathbf{MU_0} = \mathbf{\Psi}_{MU_0}\mathbf{\Gamma}^T + \mathbf{E}_{MU_0} \\ \mathbf{MU_1} = \mathbf{\Psi}_{MU_1}\mathbf{\Gamma}^T + \mathbf{E}_{MU_1} \\ \mathbf{MU_2} = \mathbf{\Psi}_{MU_2}\mathbf{\Gamma}^T + \mathbf{E}_{MU_2} \end{cases} \tag{5.1}$$

where $\mathbf{\Psi}_{M_0}$, $\mathbf{\Psi}_{H}$, $\mathbf{\Psi}_{I}$, $\mathbf{\Psi}_{L}$, $\mathbf{\Psi}_{R}$, $\mathbf{\Psi}_{MU_0}$, $\mathbf{\Psi}_{MU_1}$, and $\mathbf{\Psi}_{MU_2}$ denote $n \times p$ matrices of *score components*, $\mathbf{\Gamma}$ is a $m \times p$ matrix representing the *component loadings* whereas $\mathbf{E}_{M_0}$, $\mathbf{E}_{H}$, $\mathbf{E}_{I}$, $\mathbf{E}_{L}$, $\mathbf{E}_{R}$, $\mathbf{E}_{MU_0}$, $\mathbf{E}_{MU_1}$ and $\mathbf{E}_{MU_2}$ are $n \times m$ matrices of *residual terms*. In general, the decomposition $\mathbf{\Psi}_X\mathbf{\Gamma}^T$ yields the best $p$-rank approximation for the original matrix $\mathbf{X}$. The loadings matrix $\mathbf{\Gamma}$ contains the coefficients which relate the original variables to the new components. From an algebraical point of view, $\mathbf{\Gamma}^T$ represents the basis of the subspaces $\mathbb{R}^p$ on which each fuzzy observation is projected. Moreover, $\mathbf{\Psi}_{M_0}$, $\mathbf{\Psi}_{H}$, $\mathbf{\Psi}_{I}$,

$\mathbf{\Psi}_L$, $\mathbf{\Psi}_R$, $\mathbf{\Psi}_{MU_0}$, $\mathbf{\Psi}_{MU_1}$, $\mathbf{\Psi}_{MU_2}$ are the matrices containing the coordinates of such projections. Note that in our model representation, the internal and external spreads together with the matrices for the membership values, have all the same underlying components structure $\mathbf{\Gamma}$ (Millsap and Meredith, 1988). In general, $\mathbf{\Gamma}$ can be understood as an intermediate representation among the midpoints in $\mathbf{M_0}$, the membership values in $\mathbf{MU_0}$, $\mathbf{MU_1}$, $\mathbf{MU_2}$, and the related left ($\mathbf{H}$, $\mathbf{L}$) and right ($\mathbf{I}$, $\mathbf{R}$) spreads, respectively. This should offer a good compromise between model flexibility and model simplicity for capturing the underlying structure of the data.

### 5.2.2 Parameters estimation

In NCFCA, each empirical observation can be considered as an object represented by a $m$-dimensional polytope in $\mathbb{R}^m$. By considering the main vertices and hedges of this object, its corresponding support is an interval in $\mathbb{R}$ (for $m = 1$), a rectangle in $\mathbb{R}^2$ (for $m = 2$), a hyper-rectangle in $\mathbb{R}^m$ (for $m \geq 2$). Figure 5.1 shows an example of a generic fuzzy object in $\mathbb{R}^2$. It is

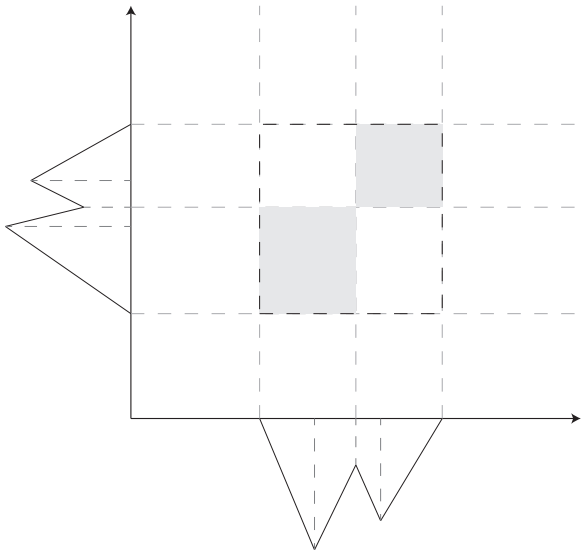


FIGURE 5.1: Example of a fuzzy object in $\mathbb{R}^2$.

interesting to note that the support of the fuzzy object is obtained by taking the union of two sub-rectangles representing the projection of the internal and external left and right spreads, respectively. Similarly, for both the fuzzy sets in the graphical representation, the projection of the modal

values $m_0$ indicates the upper and lower bounds of these sub-rectangles. Finally, the external rectangle is obtained by joining the lower and upper bounds of the left and right sub-rectangles, respectively. By using this formal framework, the parameters estimation is obtained by minimizing a suitable loss function between the observed data and the model data. To this end, several measures for fuzzy data can be adopted (Bloch, 1999). In our proposal, we resort to use a dissimilarity function based on the least squares criterion (Jahanshahloo, Lotfi, and Izadikhah, 2006; Yang and Ko, 1996):

$$\begin{aligned}\mathcal{D}^2 = &\sum_{k=1}^{2^m} \left\|(\mathbf{M_0} - \mathbf{M_0^*})\mathbf{\Phi}_k^L\right\|^2 + \\ &\sum_{k=1}^{2^m} \left\|[(\mathbf{H} - \mathbf{H}^*) + (\mathbf{L} - \mathbf{L}^*)]\mathbf{\Phi}_k^L + [(\mathbf{I} + \mathbf{I}^*) + (\mathbf{R} - \mathbf{R}^*)]\mathbf{\Phi}_k^R\right\|^2 + \\ &\sum_{k=1}^{2^m} \left\|(\mathbf{MU_1} - \mathbf{MU_1^*})\mathbf{\Phi}_k^L + (\mathbf{MU_2} - \mathbf{MU_2^*})\mathbf{\Phi}_k^R\right\|^2 + \\ &\sum_{k=1}^{2^m} \left\|(\mathbf{MU_0} - \mathbf{MU_0^*})\mathbf{\Phi}_k^L\right\|^2 .\end{aligned} \tag{5.2}$$

where $\mathbf{M_0^*} = \mathbf{\Psi}_{M_0}\mathbf{\Gamma}^T$, $\mathbf{H}^* = \mathbf{\Psi}_H\mathbf{\Gamma}^T$, $\mathbf{I}^* = \mathbf{\Psi}_I\mathbf{\Gamma}^T$, $\mathbf{L}^* = \mathbf{\Psi}_L\mathbf{\Gamma}^T$, $\mathbf{R}^* = \mathbf{\Psi}_R\mathbf{\Gamma}^T$, $\mathbf{MU_0^*} = \mathbf{\Psi}_{MU_0}\mathbf{\Gamma}^T$, $\mathbf{MU_1^*} = \mathbf{\Psi}_{MU_1}\mathbf{\Gamma}^T$, $\mathbf{MU_2^*} = \mathbf{\Psi}_{MU_2}\mathbf{\Gamma}^T$. Note that in the above function, $\mathbf{\Phi}_k^L$ and $\mathbf{\Phi}_k^R$ are $m \times m$ diagonal matrices which allow to separately consider each distinct main vertex/hedge of the $m$-dimensional polytope. More precisely, the diagonals are equal to the rows of the Boolean structural matrices $\mathbf{\Phi}^L$ and $\mathbf{\Phi}^R$ of order $2^m \times m$ which are defined according to the following properties:

(a) $\mathbf{\Phi}_t^{L/R} + \mathbf{\Phi}_{2^{(m-1)}+t}^{L/R} = \mathbf{0}_m \;\; (i = 1, \ldots, 2^{m-1})$

(b) $\mathbf{\Phi}_k^L \cdot \mathbf{\Phi}_k^R = \mathbf{0}_{m\times m}$

(c) $\mathbf{\Phi}_k^{L/R} \cdot \mathbf{\Phi}_k^{R/L} = \mathbf{\Phi}_k^{L/R}$

(d) $\mathbf{\Phi}_k^L + \mathbf{\Phi}_k^R = \mathbf{I}_{m\times m}$

(e) $\sum_{k=1}^{2^m} \mathrm{Tr}(\mathbf{X}\mathbf{\Phi}_k^{L/R}) = 2^{m-1}\,\mathrm{Tr}(\mathbf{X})$.

For instance, when $m = 2$ these matrices take the following form:

$$\mathbf{\Phi}^L = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{\Phi}^R = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Therefore, to represent the lower bounds of the support of the $m$-dimensional object, we can set $\mathbf{LB} = \mathbf{M_0} - \mathbf{H}\mathbf{\Phi}_1^L - \mathbf{L}\mathbf{\Phi}_1^L + \mathbf{I}\mathbf{\Phi}_1^R + \mathbf{R}\mathbf{\Phi}_1^R$ which, in turn, is equivalent to write $\mathbf{LB} = \mathbf{M_0} - \mathbf{H}\mathbf{\Phi}_1^L - \mathbf{L}\mathbf{\Phi}_1^L$ after noticing that $\mathbf{I}\mathbf{\Phi}_1^R = \mathbf{0}$ and $\mathbf{R}\mathbf{\Phi}_1^R = \mathbf{0}$.

The dissimilarity measure in (5.2) can also be simplified by expanding the $k-$th term of the norms by the properties (b), (c) and (e):

$$\begin{aligned} \mathcal{D}^2 = {} & 2^{m-1} \left\|(\mathbf{M_0} - \mathbf{\Psi}_{M_0}\mathbf{\Gamma}^T)\right\|^2 + 2^{m-1} \left\|(\mathbf{H} - \mathbf{\Psi}_H\mathbf{\Gamma}^T)\right\|^2 + 2^{m-1} \left\|(\mathbf{I} - \mathbf{\Psi}_I\mathbf{\Gamma}^T)\right\|^2 \\ & + 2^{m-1} \left\|(\mathbf{L} - \mathbf{\Psi}_L\mathbf{\Gamma}^T)\right\|^2 + 2^{m-1} \left\|(\mathbf{R} - \mathbf{\Psi}_R\mathbf{\Gamma}^T)\right\|^2 + 2^{m-1} \left\|(\mathbf{MU_0} - \mathbf{\Psi}_{MU_0}\mathbf{\Gamma}^T)\right\|^2 \\ & + 2^{m-1} \left\|(\mathbf{MU_1} - \mathbf{\Psi}_{MU_1}\mathbf{\Gamma}^T)\right\|^2 + 2^{m-1} \left\|(\mathbf{MU_2} - \mathbf{\Psi}_{MU_2}\mathbf{\Gamma}^T)\right\|^2 \\ & + 2^m \ \mathrm{Tr}[(\mathbf{H} - \mathbf{\Psi}_H\mathbf{\Gamma}^T)^T(\mathbf{L} - \mathbf{\Psi}_L\mathbf{\Gamma}^T)] + 2^m \ \mathrm{Tr}[(\mathbf{I} - \mathbf{\Psi}_I\mathbf{\Gamma}^T)^T(\mathbf{R} - \mathbf{\Psi}_R\mathbf{\Gamma}^T)] \end{aligned} \tag{5.3}$$

where the structural matrices are simply replaced by appropriate weights. We use the Alternating Least Squares algorithm - ALS (Kiers and Berge, 1989; Kiers, 2002) to estimate the parameters contained in $\mathbf{\Psi}_{M_0}$, $\mathbf{\Psi}_H$, $\mathbf{\Psi}_I$, $\mathbf{\Psi}_L$, $\mathbf{\Psi}_R$, $\mathbf{\Psi}_{MU_0}$, $\mathbf{\Psi}_{MU_1}$, $\mathbf{\Psi}_{MU_2}$, $\mathbf{\Gamma}$.

In particular, the final ALS solutions for the model (5.1) are:

$$\begin{aligned} \mathrm{vec}(\widehat{\mathbf{\Gamma}}) = {} & [\ (\mathbf{\Psi}_H^T\mathbf{\Psi}_H \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_L^T\mathbf{\Psi}_L \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_I^T\mathbf{\Psi}_I \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_R^T\mathbf{\Psi}_R \otimes \mathbf{I}_{m\times m}) + \\ & + (\mathbf{\Psi}_{M_0}^T\mathbf{\Psi}_{M_0} \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_{MU_0}^T\mathbf{\Psi}_{MU_0} \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_{MU_1}^T\mathbf{\Psi}_{MU_1} \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_{MU_2}^T\mathbf{\Psi}_{MU_2} \otimes \mathbf{I}_{m\times m}) + \\ & + 2(\mathbf{\Psi}_H^T\mathbf{\Psi}_L \otimes \mathbf{I}_{m\times m} + \mathbf{\Psi}_I^T\mathbf{\Psi}_R \otimes \mathbf{I}_{m\times m})\ ]^{-1} \cdot \mathrm{vec}[\ (\mathbf{H}^T\mathbf{\Psi}_H + \mathbf{L}^T\mathbf{\Psi}_L + \mathbf{I}^T\mathbf{\Psi}_I + \mathbf{R}^T\mathbf{\Psi}_R) + \\ & + (\mathbf{M_0}^T\mathbf{\Psi}_{M_0} + \mathbf{MU_0}^T\mathbf{\Psi}_{MU_0} + \mathbf{MU_1}^T\mathbf{\Psi}_{MU_1} + \mathbf{MU_2}^T\mathbf{\Psi}_{MU_2}) + \\ & + (\mathbf{H}^T\mathbf{\Psi}_L + \mathbf{L}^T\mathbf{\Psi}_H + \mathbf{I}^T\mathbf{\Psi}_R + \mathbf{R}^T\mathbf{\Psi}_I)\ ]; \end{aligned} \tag{5.4}$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_L) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T \mathrm{vec}(\mathbf{H} - \mathbf{\Psi}_H\mathbf{\Gamma}^T + \mathbf{L}); \tag{5.5}$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_R) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T \mathrm{vec}(\mathbf{I} - \mathbf{\Psi}_I\mathbf{\Gamma}^T + \mathbf{R}); \tag{5.6}$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_I) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T \mathrm{vec}(\mathbf{R} - \mathbf{\Psi}_R\mathbf{\Gamma}^T + \mathbf{I}); \tag{5.7}$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_H) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T \mathrm{vec}(\mathbf{L} - \mathbf{\Psi}_L\mathbf{\Gamma}^T + \mathbf{H}); \tag{5.8}$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_{M_0}) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T \mathrm{vec}(\mathbf{M_0}); \tag{5.9}$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_{MU_0}) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T \mathrm{vec}(\mathbf{MU_0}); \tag{5.10}$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_{MU_1}) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T \mathrm{vec}(\mathbf{MU_1}); \tag{5.11}$$

$$\mathrm{vec}(\widehat{\mathbf{\Psi}}_{MU_2}) = (\mathbf{\Gamma}^T\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^{-1} \cdot (\mathbf{\Gamma} \otimes \mathbf{I}_{n\times n})^T \mathrm{vec}(\mathbf{MU_2}); \tag{5.12}$$

where $\mathrm{vec}(.)$ is the linear operator that converts a $n \times m$ matrix into a $mn \times 1$ vector, $\otimes$ denotes the Kronecker product whereas $\mathbf{I}$ is an identity matrix of appropriate order.

In fitting the unconstrained NFCA model, we adopted an iterative procedure based on standard stopping criteria and random initialization. However, one potential limitation of such algorithm is that, in some circumstances, it might not yield feasible solutions. In particular, if the model is fitted to empirical data which are largely corrupted by noise, the corresponding estimations might violate the natural constraints of the 2-mode fuzzy numbers (namely: $\mathbf{h}^*{}_j > \mathbf{0}_n$, $\mathbf{i}^*{}_j > \mathbf{0}_n$, $\mathbf{l}^*{}_j > \mathbf{0}_n$, $\mathbf{r}^*{}_j > \mathbf{0}_n$, $\mu^*_{\mathbf{1}j} < \mu^*_{\mathbf{0}j} < \mu^*_{\mathbf{2}j}$). In these situations, a constrained version of the algorithm based on specific optimization techniques should instead be preferred (Giordani and Kiers, 2007). However, a common and easy strategy to deal with eventual infeasible parameter estimates is to apply a post-hoc correction on the estimated parameters (Giordani and Kiers, 2004a). In particular, eventual negative spreads estimates could be set to zero whereas the estimated membership values of the fuzzy data could be row-wise normalized in order to satisfy their natural constraints.

### 5.2.3 Data interpretation and visualization

Once the estimated components are finally obtained, the results can be analysed by inspecting the loadings matrix $\widehat{\mathbf{\Gamma}}$ and/or by displaying the scores in a low-dimensional plot. In particular, the loadings can be understood as linear coefficients which express the magnitudo of the relation between the observed variables and the estimated components. By contrast, the scores represent the projections of the fuzzy observation into the subspace spanned by $\widehat{\mathbf{\Gamma}}$. Like for traditional PCA (or CA), also for NCFCA the score plot represents an important visualization procedure that allows to assess the relationship among the projected units (e.g., by analysing the patterns of similarity or dissimilarity among the units). In what follows, we describe in more details both the data pre-treatment technique and the data evaluation procedure adopted in NCFCA modelling.

#### Data pre-treatment

A common practice in multivariate analysis is to pre-process raw-data in order to obtain an improved and clean dataset. Two of the most important ways to pre-process raw-data are *centering* and *scaling*. Centering corresponds to a repositioning of the coordinate system such that the center of gravity of the cloud of data points becomes the origin. By contrast, scaling allows to re-distribute the data according to a specific factor (e.g., the standard deviation). In particular, scaling configures the original variables within a unique scale range without changing the original structure of the data. Centering and scaling can be performed for several reasons. For instance, centering may be applied in order to improve the fit of the model, remove noise from the data, avoid problems in the estimation procedures, etc. Similarly, scaling may be implemented in order to adjust for scale differences, reduce the inflation of small or big values in the data, improve the interpretation and visualization of the results, etc. (Bro and Smilde, 2003; Berg et al., 2006). Several methods can be used for centering and scaling (e.g., auto-scaling, range-scaling, pareto-scaling, vast-scaling,

etc). In this contribution, we opted for mean-centering and pareto-scaling methods. In general, given a matrix $\mathbf{A}$, the mean-centered matrix $\mathbf{A}^*$ is obtained as $\mathbf{A}^* = \mathbf{A} - \mathbf{1}_{n\times m} \cdot \mathrm{diag}(\bar{\mathbf{a}})$ whereas the pareto-scaling is performed by $\mathbf{A}^* = \mathbf{A} \cdot [\mathrm{diag}(\sqrt{\mathrm{std}(\mathbf{A})})]^{-1}$. Note that, $\bar{\mathbf{a}}$ is the vector containing the column means of $\mathbf{A}$, $\mathrm{std}(.)$ indicates the column-wise standard deviation whereas $\mathrm{diag}(.)$ is the operator which transforms a vector into a diagonal matrix. In particular, we pre-processed our data matrices according to the following steps: (i) $\mathbf{M_0}$ was simultaneously mean-centered and pareto-scaled by considering its means and standard deviations, (ii) $\mathbf{H}$ and $\mathbf{I}$ were pareto-scaled by considering the standard deviations of $\mathbf{M_0}$, (iii) $\mathbf{L}$ and $\mathbf{R}$ were pareto-scaled by considering the standard deviations of $\mathbf{M_1}$ and $\mathbf{M_2}$ respectively, (iv) $\mathbf{MU_0}$, $\mathbf{MU_1}$ and $\mathbf{MU_2}$ were pareto-scaled by considering the standard deviations of $\mathbf{M_0}$, $\mathbf{M_1}$ and $\mathbf{M_2}$, respectively.

**Rotation of $\widehat{\mathbf{\Gamma}}$**

Unlike standard PCA (or CA), the NCFCA estimation procedure does not necessarily yield an orthonormal matrix $\widehat{\mathbf{\Gamma}}$. For this reason, a direct interpretation of $\widehat{\mathbf{\Gamma}}$ might be arduous for some datasets. However, by adopting an orthonormalization procedure such as, for example, the *modified Gram-Schmidt* algorithm, one can always define a rotation matrix $\mathbf{\Omega}$ such that $\widehat{\mathbf{\Gamma}}\mathbf{\Omega}$ is column-wise orthonormal (Trefethen and Bau, 1997). In particular, the modified Gram-Schmidt algorithm requires to balance the estimated score matrices with the inverse of the transpose of $\mathbf{\Omega}$. For instance, by considering the case of $\mathbf{M_0}$ the balancing is performed as $\widehat{\mathbf{\Psi}}_{M_0}(\mathbf{\Omega}^T)^{-1}$. In addition, in order to facilitate the interpretation of the components structure, the analysis might also involve a rotation of $\widehat{\mathbf{\Gamma}}\mathbf{\Omega}$. Several techniques can be adopted to this purpose (Kiers, 1997). In NCFCA modelling we adopted the well-known *Varimax rotation* which provides a very simple components structure where each original fuzzy variable is associated with a small set of components (Kaiser, 1958). The rotation of $\widehat{\mathbf{\Gamma}}$ allows to simplify the interpretation of both numerical and graphical results of the model.

### Model evaluation

In this subsection we illustrate some useful procedures to assess the performance and reliability of the NCFCA model.

*Goodness of fit*. In order to evaluate the performance of the NCFCA model, we considered the normalized index:

$$R = 1 - (A/B)$$

where:

$$\begin{aligned} A = &\left\|\mathbf{M_0} - \widehat{\mathbf{\Psi}}_{M_0}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \left\|\mathbf{H} - \widehat{\mathbf{\Psi}}_{H}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \left\|\mathbf{I} - \widehat{\mathbf{\Psi}}_{I}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \left\|\mathbf{L} - \widehat{\mathbf{\Psi}}_{L}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \left\|\mathbf{R} - \widehat{\mathbf{\Psi}}_{R}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \\ &+ \left\|\mathbf{MU_0} - \widehat{\mathbf{\Psi}}_{MU_0}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \left\|\mathbf{MU_1} - \widehat{\mathbf{\Psi}}_{MU_1}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \left\|\mathbf{MU_2} - \widehat{\mathbf{\Psi}}_{MU_2}\widehat{\mathbf{\Gamma}}^T\right\|^2 + \\ &+ \mathrm{Tr}[(\mathbf{H} - \widehat{\mathbf{\Psi}}_{H}\widehat{\mathbf{\Gamma}}^T)^T(\mathbf{L} - \widehat{\mathbf{\Psi}}_{L}\widehat{\mathbf{\Gamma}}^T)] + \mathrm{Tr}[(\mathbf{I} - \widehat{\mathbf{\Psi}}_{I}\widehat{\mathbf{\Gamma}}^T)^T(\mathbf{R} - \widehat{\mathbf{\Psi}}_{R}\widehat{\mathbf{\Gamma}}^T)] \end{aligned}$$

$$\begin{aligned} B = &\|\mathbf{M_0}\|^2 + \|\mathbf{H}\|^2 + \|\mathbf{I}\|^2 + \|\mathbf{L}\|^2 + \|\mathbf{R}\|^2 + \|\mathbf{MU_0}\|^2 + \|\mathbf{MU_1}\|^2 + \|\mathbf{MU_2}\|^2 + \\ &+ \mathrm{Tr}[\mathbf{H}^T\mathbf{L}] + \mathrm{Tr}[\mathbf{I}^T\mathbf{R}]. \end{aligned}$$

The index $R$ takes values in $[0, 1]$ and compares the residual sum of squares ($A$) with the observed sum of squares ($B$). High values for this index indicate that the model appropriately fits the data (Giordani, 2010; Bro and Smilde, 2003).

*Reliability*. To asses the accuracy of the NCFCA solutions, we used a non-parametric bootstrap procedure for component analysis (Coppi, Giordani, and D'Urso, 2006; Kiers, 2004). In particular, in the non-parametric bootstrap $Q$ samples (with $Q \geq 1000$) of size $n$ were row-wise randomly drawn (with replacement) from the original matrices $\mathbf{M_0}$, $\mathbf{H}$, $\mathbf{I}$, $\mathbf{L}$, $\mathbf{R}$, $\mathbf{MU_0}$, $\mathbf{MU_1}$, $\mathbf{MU_2}$. For each $q$-th sample, the loadings matrix $\widehat{\mathbf{\Gamma}}^q$ was derived by applying the NCFCA procedure on the sample matrices $\mathbf{M_0}^q$, $\mathbf{H}^q$, $\mathbf{I}^q$, $\mathbf{L}^q$, $\mathbf{R}^q$, $\mathbf{MU_0}^q$, $\mathbf{MU_1}^q$, $\mathbf{MU_2}^q$. In order to make bootstrap solutions optimally comparable, $\widehat{\mathbf{\Gamma}}^q$ was rotated to match as close as possible the original $\widehat{\mathbf{\Gamma}}$. Such rotation was obtained by finding a rotation matrix $\mathbf{\Omega}^q$ that minimizes the risk $\|\widehat{\mathbf{\Gamma}}^q\mathbf{\Omega}^q - \widehat{\mathbf{\Gamma}}\|^2$ with the following optimal solution $\mathbf{\Omega}^q = (\widehat{\mathbf{\Gamma}}^{qT}\widehat{\mathbf{\Gamma}}^q)^{-1}\widehat{\mathbf{\Gamma}}^{qT}\widehat{\mathbf{\Gamma}}$. These steps were then repeated for $Q$ times. Finally, the ensuing rotate

sample matrices $\widehat{\mathbf{\Gamma}}^1, \widehat{\mathbf{\Gamma}}^2, \ldots, \widehat{\mathbf{\Gamma}}^Q$ were used for computing (by re-sampling) the standard errors or percentile intervals for every estimated parameter in the model. In general, the lower the standard errors, the greater the accuracy of the model.

### Score plot

The score plot is the graphical representation of the original $n$ units in the $\mathbb{R}^p$ subspace. Unlike standard PCA (or CA), in the NCFCA framework each statistical unit is represented by an hyper-rectangle in $\mathbb{R}^p$. There are several methods that can be considered for score plotting, such as for example Maximum Covering Area Rectangle - MCAR (Cazes et al., 1997), Parallel Edge Connected Shapes - PECS (Irpino, Lauro, and Verde, 2003) and Polytope Representation (Le-Rademacher and Billard, 2012). In this contribution we adopted the MCAR approach which is a simple and fast graphical technique to represent interval, symbolic or fuzzy data. In particular, in the simple two dimensional case, MCAR allows to illustrate the statistical units by means of rectangles in $\mathbb{R}^2$ whereas the information associated to the membership functions is usually not represented.[1] More formally, in our context MCAR was applied as follows. Once the loadings matrix was columnwise orthonormalized and the score matrices balanced, each non-convex fuzzy data was described as the union of two rectangles (or hyper-rectangles) referring to the left and right internal and external spreads (see Figure 5.1). The vertices representing the lower and upper bounds of the external rectangles were obtained by the score matrices using the following formula:

$$\begin{aligned}\mathbf{\Lambda}_i = \mathbf{\Phi}^L_{2^p\times p}[\ \mathrm{diag}(\widehat{\mathbf{\Psi}}_{M_0 i}) - \kappa|\ \mathrm{diag}(\widehat{\mathbf{\Psi}}_{Hi})| - \kappa|\ \mathrm{diag}(\widehat{\mathbf{\Psi}}_{Li})|\ ]+ \\ + \mathbf{\Phi}^R_{2^p\times p}[\ \mathrm{diag}(\widehat{\mathbf{\Psi}}_{M_0 i}) + \kappa|\ \mathrm{diag}(\widehat{\mathbf{\Psi}}_{Ii})| + \kappa|\ \mathrm{diag}(\widehat{\mathbf{\Psi}}_{Ri})|\ ]\end{aligned} \tag{5.13}$$

[1] Note, however, that the membership functions always contribute to the orientation of the axes in $\mathbb{R}^p$ even if they are not directly illustrated in the graphical representation.

whereas the inner left and right rectangles were depicted by considering the midpoints in $\widehat{\mathbf{\Psi}}_{M_0}$ as upper and lower vertices, respectively. Note that $\mathbf{\Phi}^L_{2^p \times p}$ and $\mathbf{\Phi}^R_{2^p \times p}$ are Boolean structural matrices of order $2^p \times p$ having the same structures and properties of those described in Section 5.2, whereas $|.|$ indicates the absolute value. Note that in (5.13) the scalar $\kappa \in ]0, 1]$ acts as a resizing factor which allows to reduce the eventual oversize-effect of the plotted rectangles.

## 5.3 Applications

In this section we describe three applications to illustrate the main features of the NCFCA analysis.

### 5.3.1 Example 1: psychological assessment of worry

In this first example we analysed a real dataset about the psychology of worry (Stöber and Joormann, 2001). In clinical psychology, the assessment of worry is usually characterized by high levels of imprecision and vagueness in the data. The clinical inventory was composed by 7 items (see Table 5.1) and administered to a group of 10 undergraduate students from the University of Trento (Italy). The scores were collected using a computerized interface based on the mouse tracking methodology (Calcagnì, Lombardi, and Pascali, 2013; Johnson et al., 2012). In particular, for each of the seven items, participants were told that a pseudo-circular scale with five response levels (strongly disagree, disagree, neither, agree, strongly agree) would be presented on the screen, and that they were asked to choose which of these responses was the most appropriate for the presented item. After participants clicked a start button, a window with the text of the item appeared at the top of the screen. Next the scale with the five levels appeared while the cursor was allocated to the center of the screen. Participants give their responses by mouse-clicking the chosen level of the scale. Meanwhile, we recorded the streaming x-y coordinates of the computer mouse. Figure 5.2-a/b shows two empirical patterns of

mouse movements. In particular, figure 5.2-a represents an empirical pat-

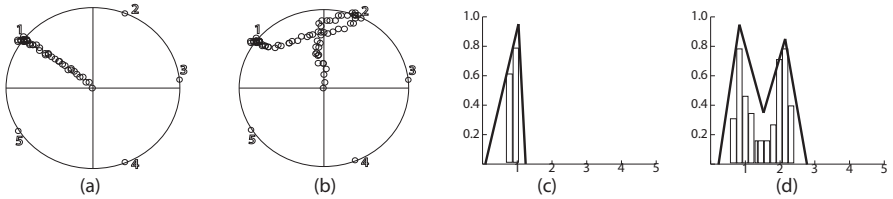



FIGURE 5.2: Mouse Tracker: empirical patterns of mouse movements (a/b) and their associated histograms and fuzzy sets (c/d). Note that: the two patterns (a) and (b) are different while sharing a same finale response (1 = strongly disagree), the numbers encode the five levels of the scale (1 = strongly disagree, 2 = disagree, 3 = neither, 4 = agree, 5 = strongly agree), whereas the histograms were rescaled in order to provide a better comparison with the fuzzy sets.

tern with a low imprecision/fuzziness, by contrast figure 5.2-b shows a pattern with a higher level of imprecision and vagueness. Figure 5.2-c/d shows the histograms and the associated fuzzy sets constructed using the radial positions of the x-y mouse movement coordinates of the empirical patterns. In particular, the fuzzy sets were obtained by a heuristic optimization procedure that allowed to convert histograms into fuzzy sets (Ciavolino, Salvatore, and Calcagnì, 2013). Before running the NCFCA analysis, the datasets were first pre-processed according to the procedure described in Sec. 5.5.1, next the NCFCA algorithm was used to extract two main components ($p = 2$). The algorithm converged after only 30 iterations and the overall goodness of fit of the model was good ($R = 0.90$). Moreover, model accuracy and reliability was also good as indicated by the low standard errors reported in Table 5.1. Finally, the estimated loadings matrix was orthonormalized and varimax-rotated to simplify the interpretation of the components in the NCFCA model. Figure 5.3 shows an example of some observed and reconstructed fuzzy sets on the fifth variable ($x_5$).

In order to identify the meaning of each extracted component, we selected the variables with loading values larger than $\pm 0.35$ (relevant variables). The results reported in Table 5.1 showed that the first component depended on $x_2$ (-0.50) and $x_5$ (-0.43) and, to a less extent, on $x_1$ (-0.38), $x_6$ (-0.39) and $x_7$ (-0.39), whereas the second component exclusively depended

| Clinical items/Variables | Comp. 1 | Comp. 2 |
|---|---|---|
| ($x_1$) I am good at controlling negative and positive emotions | **-0.38**(0.03) | 0.00(0.05) |
| ($x_2$) I am worried that I will never realize my ambitions | **-0.50**(0.04) | 0.10 (0.08) |
| ($x_3$) It is important that human relations are based upon trust | -0.01(0.04) | **-0.98**(0.16) |
| ($x_4$) It is important that I am competent in everything I do | -0.34(0.04) | -0.10(0.08) |
| ($x_5$) I am worried to be bad | **-0.43**(0.04) | -0.01(0.05) |
| ($x_6$) I am worried to lose my close friends | **-0.39**(0.05) | 0.09(0.10) |
| ($x_7$) I like to be alone when I am working with a problem | **-0.39**(0.03) | -0.10(0.04) |

TABLE 5.1: Example 1: Loadings matrix $\widehat{\mathbf{\Gamma}}$ with standard errors in parenthesis ($Q = 5000$). Loadings higher than 0.35 (in absolute sense) are in boldface type.

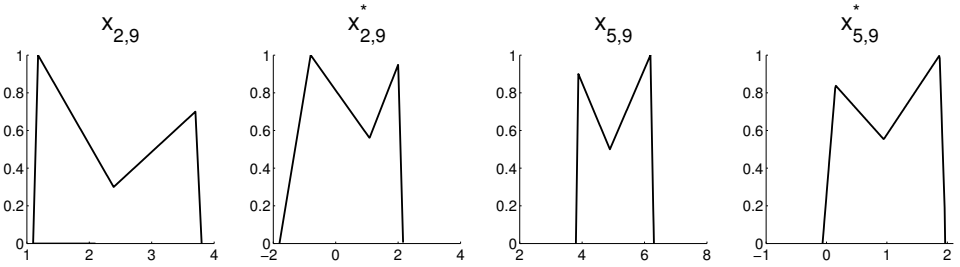



FIGURE 5.3: Example 1: Observed ($x$) and model reconstructed ($x^*$) fuzzy sets for the subjects 2 and 9 on $x_5$

on $x_3$ (-0.98). Therefore, taking into account the meaning of the significant variables, the first component can be understood as 'individual dimension' whereas the second component can be referred as 'interpersonal dimension' of psychology of worry.

Figure 5.4 shows the score plot for the NCFCA model in the two dimensional space spanned by $\widehat{\mathbf{\Gamma}}$. A substantive interpretation of Figure 5.4 can be provided by considering both the positions and sizes of the rectangles. In particular, the size of the rectangles reflected the imprecision associated with the clinical items which played a significant role in the definition of the two components. In our case, the fuzzy units were arranged into three main regions: central (units 1, 4, 5, 6, 7, 10), left-outer (units 2, 9), right-outer (units 3, 8). In particular, the individuals in the middle part of the plot showed clinical scores that were in the mean range for both the components. It is important to note that due to the mean-centering procedure, the origin of the axis represents the *average region* and therefore the units located in this area are characterized by *mean profiles*. On the contrary, the units located far away from the center of gravity of the plot shows typical

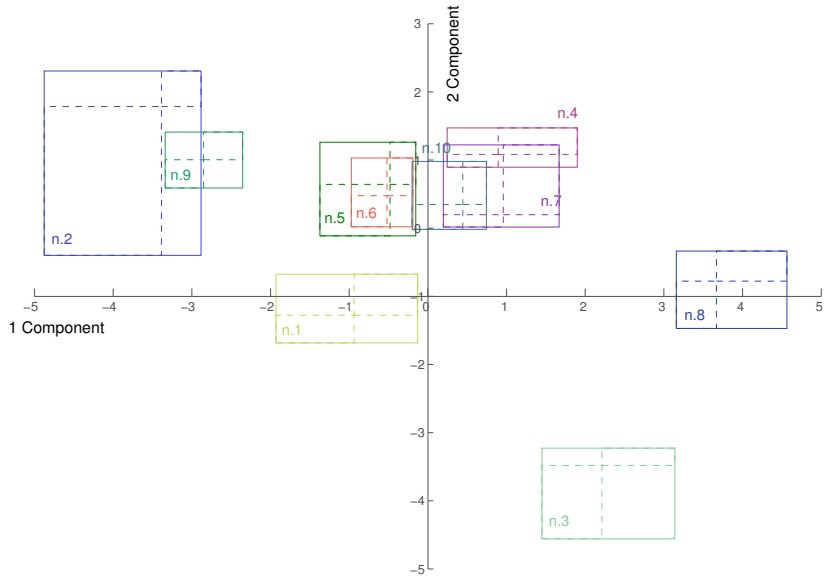



FIGURE 5.4: Example 1: Score plot for the first and second components (fuzzy units are consecutively numbered and represented with different colours)

features which did not belong to the mean profile.

### 5.3.2 Example 2: Self perception of professional roles

In this second application we studied a real dataset about psychologists' self perception of their professional role. To this end, a specific 7-items questionnaire (see Table 5.2) was administered to a group of 24 psychology students from the University of Trento (Italy). Data were collected by means of a computerized questionnaire based on fuzzy rating scales commonly used in human ratings studies (Hesketh, Pryor, and Hesketh, 1988). In particular, the fuzzy scale was based on a pseudo-continuous scale implemented using a suitable graphical interface (see Figure 5.5). Interestingly, the fuzzy rating scale may elicit to different scenarios: the respondent chooses a single level of the scale (Figure 5.5-b) or s(he) selects an intermediate position which lies between the two levels (Figure 5.5-c). The NCFCA model was applied to the pre-processed data and three components ($p = 3$) were extracted using the NCFCA algorithm. The algorithm converged after 112 iterations and the goodness-of-fit of the model was good ($R = 0.88$). Model accuracy and reliability were also good as

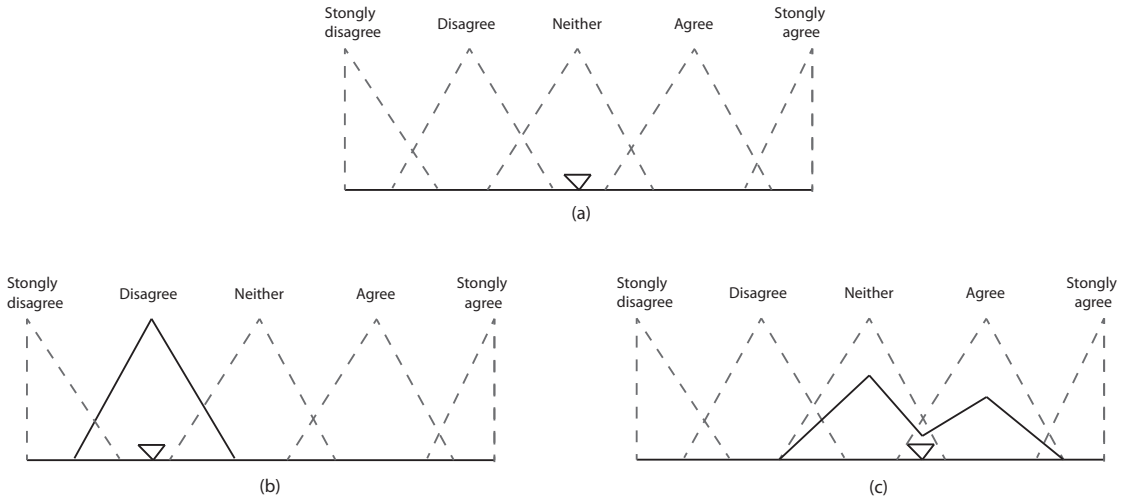



FIGURE 5.5: Fuzzy rating scale with a pseudo-continuous line, a movable cursor (in solid line) and a fuzzy variable (in dotted line) representing the hidden fuzzy levels of the scale.

shown by the low standard errors reported in Table 5.2. Like for the previous analysis, also in this second application the estimated loading matrix was orthonormalized and varimax-rotated in the NCFCA model.

| Variables | Comp. 1 | Comp. 2 | Comp. 3 |
|---|---|---|---|
| ($x_1$) Degree of development of Italian university system | **-0.46**(0.07) | 0.03(0.06) | -0.25(0.08) |
| ($x_2$) Usefulness of university studies | 0.23(0.06) | -0.17(0.07) | **-0.80**(0.10) |
| ($x_3$) Trust in local government | -0.26(0.06) | 0.17(0.07) | **-0.53**(0.08) |
| ($x_4$) Trust in private enterprise | **-0.46**(0.07) | -0.20(0.06) | 0.03(0.08) |
| ($x_5$) Degree of devel. of Italy | **-0.67**(0.06) | 0.04(0.06) | 0.09(0.07) |
| ($x_6$) Psychologist is necessary | -0.08(0.06) | **-0.53**(0.05) | -0.05(0.06) |
| ($x_7$) Psychologist is useful | 0.02(0.07) | **-0.79**(0.06) | 0.08(0.09) |

TABLE 5.2: Example 2: Loadings matrix $\widehat{\mathbf{\Gamma}}$ with standard errors in parenthesis ($Q = 5000$). Loadings higher than 0.35 (in absolute sense) are in boldface type.

The results reported in Table 5.2 showed that the first component was inversely related to $x_1$ (-0.46), $x_4$ (-0.46) and $x_5$ (-0.67). Similarly, the second component was also inversely related to $x_6$ (-0.53) and $x_7$ (-0.79). Finally, the third component inversely depended on $x_2$ (-0.80) and $x_3$ (-0.53). In line with these results, the first component can be interpreted as 'future perspective dimension', the second component refers to 'psychology as profession', whereas the third component can be understood as 'present dimension'. Figure 5.3 shows an example of some observed and reconstructed fuzzy sets on the second variable ($x_2$). Figures 5.7 and 5.8 show the score plots for the first vs. second and second vs. third components,

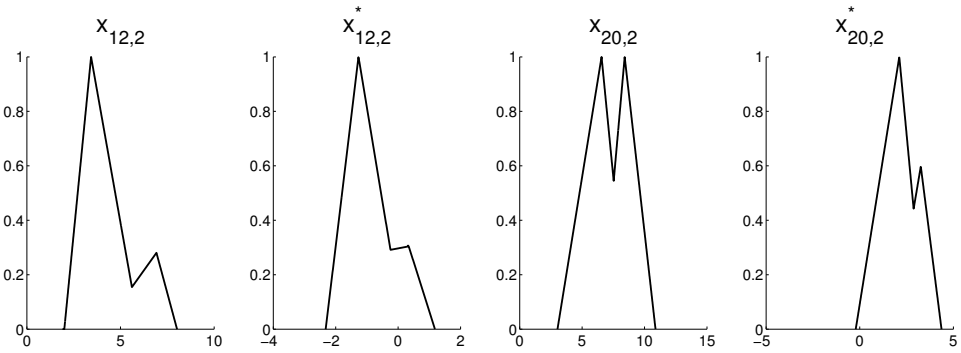



FIGURE 5.6: Example 2: Observed ($x$) and model reconstructed ($x^*$) fuzzy sets for the subjects 12 and 20 on $x_2$

respectively. In particular, the first score plot (Figure 5.7) contrasts *future*

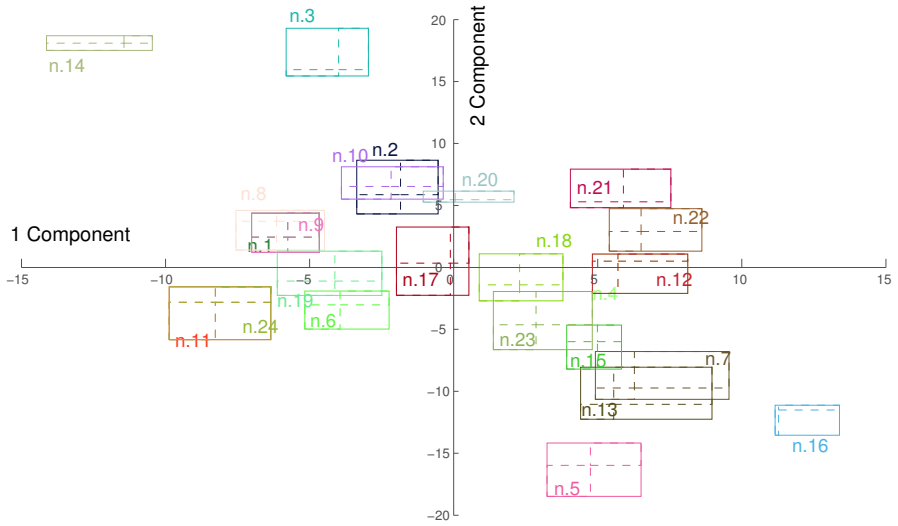



FIGURE 5.7: Example 2: Score plot for the first and second components (fuzzy units are consecutively numbered and represented with different colours)

and *profession*. It shows an interesting pattern in which most of the units were located on the middle part of the plot whereas only small groups of units are located in the left-outer (units 3,14) and right-outer (units 5,7,13,16) parts of the graphical representation. The second score plot (Figure 5.8) contrasts *present* and *profession*. Like for the previous graphical representation, also for the second score plot most of the units are located in the middle portion of the plot. Finally, the sizes of the rectangles were small and similar among dimensions (in general, the students seemed to convey the same degree of imprecision in providing their responses).

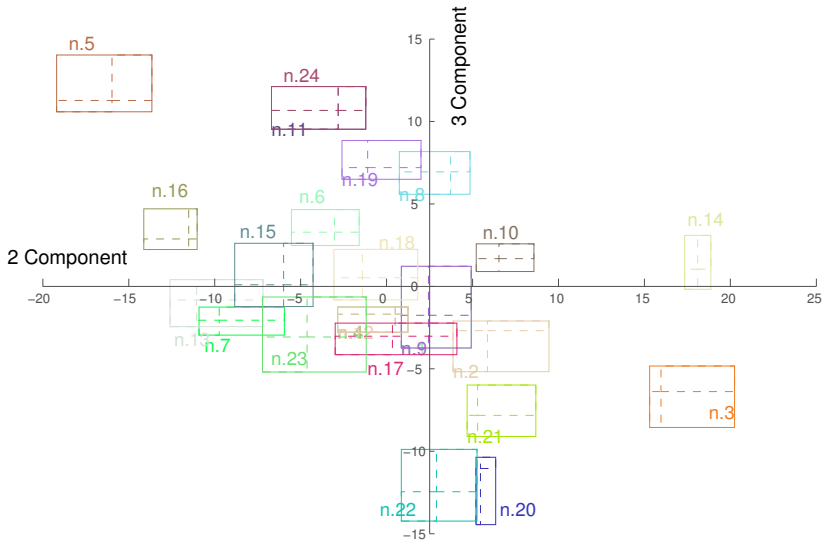



FIGURE 5.8: Example 2: Score plot for the second and third components (fuzzy units are consecutively numbered and represented with different colours)

### 5.3.3 Example 3: welfare and productivity of Italian regions

In this last example we tested our model on a real dataset about economic and social indicators collected by the National Institute of Statistics (ISTAT). The original dataset contained 10 socio-economic indicators referred to 20 Italian regions (see table 5.3). In order to test our model the crisp variables were first re-scaled to a common scale and next fuzzified with a suitable fuzzification procedure based on the Mamdani fuzzy system (Lalla, Facchinetti, and Mastroleo, 2005). The fuzzification routine yielded the following sets for the fuzzy variables: null (0,0,1.67), very-low (0,2.13,3.33), low (1.67,3.87,5.0), middle (3.33,4.69,6.67), high (5.0,7.47,8.33), very-high (6.67,8.95,10.0), extreme (8.33,10.0,11.67). The NCFCA algorithm was applied to the pre-processed data and two components ($p = 2$) were extracted. The algorithm converged after 20 iterations and the goodness-of-fit of the model was generally good ($R = 0.90$). Model accuracy and reliability were also good as shown by the low standard errors reported in Table 5.3. Finally. the estimated loading matrix was orthonormalized and varimax-rotated. By inspecting Table 5.3, one can observe as the first component was negatively related to $x_1$ (-0.41), $x_2$ (-0.43), $x_3$ (-0.42), $x_4$

| Indicators/Variables | Comp. 1 | Comp. 2 |
|---|---|---|
| ($x_1$) Household spending | **-0.41**(0.01) | 0.04(0.02) |
| ($x_2$) Investments | **-0.43**(0.02) | 0.03(0.01) |
| ($x_3$) Income | **-0.42**(0.02) | 0.01(0.01) |
| ($x_4$) Salaries | **-0.40**(0.01) | 0.05(0.02) |
| ($x_5$) Marriage index | -0.04(0.02) | **0.49**(0.02) |
| ($x_6$) Public education expenditures | 0.04(0.01) | **0.54**(0.01) |
| ($x_7$) Unemployment index | 0.04(0.01) | **0.53**(0.01) |
| ($x_8$) Energy consumption | -0.05(0.02) | **0.42**(0.03) |
| ($x_9$) Public culture expenditures | **-0.37**(0.02) | -0.06(0.04) |
| ($x_{10}$) Efficiency of health index | **-0.41**(0.02) | -0.06(0.02) |

TABLE 5.3: Example 3: Loadings matrix $\widehat{\mathbf{\Gamma}}$ with standard errors in parenthesis ($Q = 5000$). Loadings higher than 0.35 (in absolute sense) are in boldface type.

(-0.40), $x_{10}$ (-0.41) and to a less extent to $x_9$ (-0.37). By constrast, the second component was positively related to $x_5$ (0.49), $x_6$ (0.54), $x_7$ (0.53) and $x_8$ (0.42). The first component can be interpreted as 'overall productivity' and the second one can be referred to 'territorial welfare'. Figure 5.10 shows an example of some observed and reconstructed fuzzy sets on the seventh variable ($x_7$). Figure 5.10 shows the score plot for the extracted

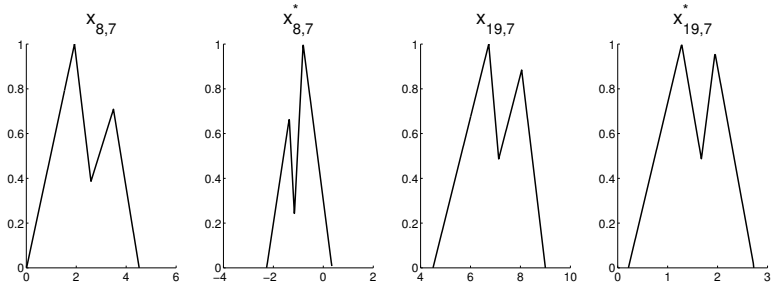



FIGURE 5.9: Example 3: Observed ($x$) and model reconstructed ($x^*$) fuzzy sets for the subjects 8 and 19 on $x_7$

components. A clear pattern can be read from the score plot. In particular, the southern regions (Campania, Calabria, Sicily, Sardinia, Apulia, Molise) were located in the second quadrant, whereas many of the richest northern regions (Lombardy, Veneto, Emilia Romagna, Tuscany) were in the fourth quadrant. Finally, the middle part of the plot contained central regions as well as some small northern ones (e.g., Friuli, Aosta Valley, Trentino). The regions in the fourth quadrant showed socio-economical profiles characterized by high productivity as well as a solid territorial welfare, whereas

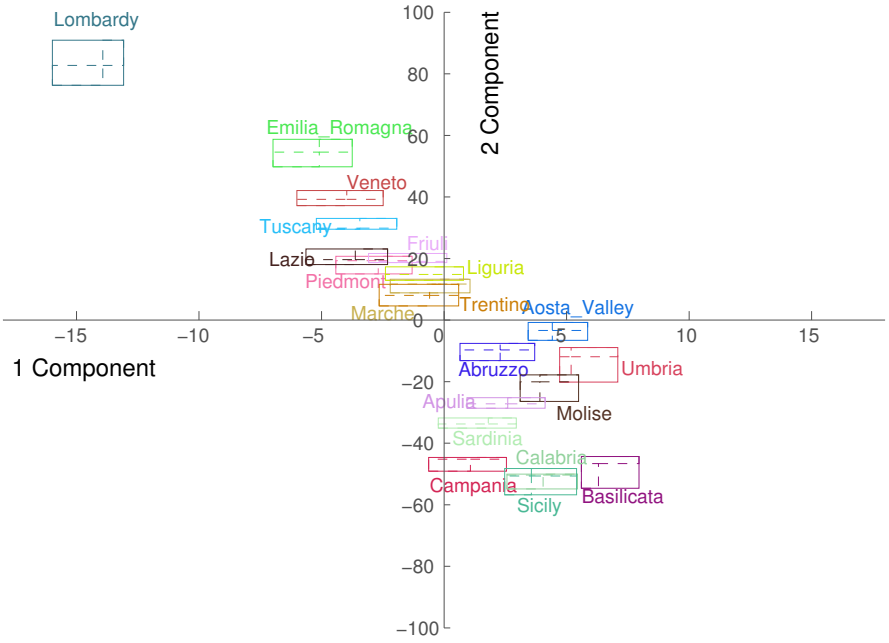



FIGURE 5.10: Example 3: Score plot for the first and second components (fuzzy units are named and represented with different colours)

the regions in the second quadrant of the map showed the opposite pattern. In particular, the regions Campania, Sicily, Calabria and Basilicata showed low values for productivity and territorial welfare. By considering the sizes of the projected rectangles, most of the regions were characterized by similar degree of fuzziness whereas Lombardy seemed to be the region with the higher fuzziness.

## 5.4 Conclusion and further perspectives

In this chapter we extended the component analysis approach to non-convex fuzzy data. The proposed NCFCA method allowed to reduce the dimensionality of multivariate datasets with non-convex fuzzy observations. In particular, the proposed method considered non-convexity by directly incorporating the membership values of fuzzy observations in the NCFCA model. To better illustrate the NCFCA features, we also described three real applications with non-convex fuzzy data. The empirical results suggested the important role played by this property when

researchers have to deal with complex, imprecise and vague information. Furthermore, it is straightforward to note how NCFCA can also be applied when data are represented by standard convex fuzzy features, in particular by setting $\mathbf{MU_0} = \mathbf{MU_1} = \mathbf{MU_2} = \mathbf{1}_{n\times m}$ for trapezoidal cases and $\mathbf{H} = \mathbf{I} = \mathbf{0}_{n\times m}$ together with $\mathbf{MU_0} = \mathbf{MU_1} = \mathbf{MU_2} = \mathbf{1}_{n\times m}$ for triangular cases.

However, the proposed method can potentially suffer from some limitations. For instance, in some empirical cases the piecewise linear representation for 2-mode fuzzy data cannot be valid and therefore other representation should be preferred (e.g., quadratic or cubic 2-mode fuzzy numbers). Moreover, for some applications the unconstrained algorithm could estimate fuzzy data which violate their normative representation and the MCAR-representation may not adequately represent the whole information provided by 2-mode fuzzy data.

Various possible extensions of our proposal could be considered. For example, the adoption of a constrained approach for the estimation procedure would always guarantee the consistency between the estimated parameters and the corresponding normative representations for non-convex fuzzy numbers. A future venue of research, would also consist in the improvement of the graphical representation as well as in the extension of the NCFCA method to cases where data are described by means of 2-mode fuzzy data beyond the ones characterized by piecewise linear membership functions. Finally, an organic framework able to deal with randomness and fuzziness simultaneously may be adopted (e.g., by using Fuzzy Random Variables). This would extend our proposal beyond the semi-descriptive approach presented in this contribution.

# Chapter 6

# Multiple mediation analysis for interval-valued data

The content of the chapter has been submitted as a research chapter for *Psychometrika* and is currently under review.

## 6.1 Introduction

In behavioral and social sciences, mediation analyses are widely used techniques for modeling underlying mechanisms of complex relationships in empirical data (MacKinnon and Fairchild, 2009; Edwards and Lambert, 2007). For instance, in working psychology the relation between job autonomy and job satisfaction/performance is better clarified by adding employees' perceived control of their time as a third variable in the original bi-variate path job autonomy-job satisfaction/performance (Claessens et al., 2004). Similarly, in clinical psychology the relation between the effect of a clinical program on adolescent antisocial behaviors is better explained by considering two intervening variables, namely the reductions in deviant peer associations and the improved family management skills (Eddy and Chamberlain, 2000). Likewise, in prevention science, beliefs, attitudes, motivations, and social support usually mediate the impact of dietary programs on the percentage of fat intake, servings of vegetables, and weekly physical activity (Kristal et al., 2000). In general, mediation analyses can be used to represent situations in which the observed relation between an independent variable, $x$, and a dependent variable, $y$, is explained using a set of third variables, $m_1, m_2, \ldots, m_k$, in the relation,

called mediator variables. Considering the cause-effect relation between $x$ and $y$, the mediator variables are causally placed between $x$ and $y$ such that a change in $x$ produces changes in $m_j$ ($j = 1 \dots k$) which, in turn, also produce changes in $y$. The $x$-$m_j$-$y$ pathway explains the process through which $x$ partially (or fully) acts on $y$ (Baron and Kenny, 1986; Yuan, Cheng, and Maxwell, 2013).

Traditionally, mediation hypotheses have been addressed and implemented within the well-known linear least squares approach (Baron and Kenny, 1986; Bollen and Stine, 1990; Edwards and Lambert, 2007; Judd and Kenny, 1981; MacKinnon, 2008). In this context, the direct and indirect pathways linking the observed variables are modeled by three or more linear equations whereas the computation of the derived model's effects is performed once the model's parameters are successfully estimated, for instance using the Baron and Kenny's steps procedure, Sobel test (Sobel, 1982), distribution of the product (Preacher, Rucker, and Hayes, 2007), and/or bootstrap based methods (Preacher and Hayes, 2008). However, over the years mediation analysis has been also addressed with the more general framework of the *causal inference analysis* (Pearl et al., 2009; Imai and Van Dyk, 2004). By using a general statistical language based on the counterfactual analysis and taking the advantages of treatment regimes and randomized control trials, this approach to mediation hypotheses allows to define general models handling with linear and non-linear relationships, parametric and non-parametric models, and continuous or discrete mediators (Imai and Van Dyk, 2004; Imai, Keele, and Tingley, 2010; Vanderweele and Vansteelandt, 2009). In such a framework, models' estimation can be performed, for instance, with sensitivity analysis (Imai, Keele, and Yamamoto, 2010), instrumental variables (Nkurunziza and Ejaz Ahmed, 2011), and principal stratification (Gallop et al., 2009).

In general, mediation models have been developed for single-valued data only. However, in some empirical contexts the observed information may

show more complex structures or patterns than those commonly represented by single-valued data. For example, when large datasets are summarized into smaller and more manageable ones, single statistical units can be described by classes of different attributes or structures (e.g., interval-valued data, histogram-valued data, symbolic-valued data) (Billard and Diday, 2003). In particular, *interval-valued data* can be understood as the simplest and most widely known type of structured data. In particular, interval-valued data may potentially arise in different empirical contexts such as, for instance, when (i) three-way datasets are reduced in two-way structures (Diday, Noirhomme-Fraiture, et al., 2008), (ii) clinical patient course and/or repeated measures are summarized by adopting procedures like the *response feature analysis* (Frison and Pocock, 1992; Everitt, 1995; Arndt et al., 2000), (iii) confidential data are masked by summarized data (Little, 1993), (iv) empirical data are modeled by interval semi-orders (Luce, 1956; Fishburn, 1973; Halff, Ortony, and Anderson, 1976), (v) observed measures are affected by systematic uncertainty (Dai, Wang, and Mi, 2013; Salicone, 2007; Rowe, 1994).

In this contribution we propose a tailor-made approach, named *interval mediation analysis* (IMedA), for multiple mediation analysis with interval valued data. This framework is based on the methodological procedures addressing mediation in the context of least squares regression and path analysis (Edwards and Lambert, 2007) and the symbolic regression for interval-valued data (Billard and Diday, 2002; Lima Neto and Carvalho, 2008). Despite its simplicity, the regression based mediation remains the most simple and widely adopted approach in conducting mediation analysis. It also guarantees that the extension of the mediation analysis to the interval-valued data still remains at a manageable level of technical complexity.

The remainder of the chapter is organized as follows. Section 6.2 briefly recalls the basic characteristics of interval-valued data and their application in psychological research. Section 6.3 exposes the proposed multiple

mediation model for interval-valued data together with their main properties. Sections 6.4 and 6.5 illustrate procedures for data analysis and model evaluation. Section 6.6 reports a brief Monte Carlo simulation study to evaluate the performance of the IMedA algorithm. Section 6.7 describes two case studies showing the application of the new approach to some psychological datasets. Finally, Section 6.8 concludes this chapter providing final remarks and suggestions for future extensions of the current contribution.

## 6.2 Interval-valued data

Here we describe the concept of interval-valued data, their properties, and some examples of application in psychological research. Note that although different types of interval representations have been defined in the mathematical/statistical literature (e.g., open, closed, half-open, bounded, and unbounded intervals), in this article we will limit our attention to *closed bounded intervals*, only.

### 6.2.1 Introduction and examples of application

Interval data are structured data that can be applied in modeling several empirical situations where the knowledge to be extracted is complex and/or highly structured. Unlike single-valued data, which can just represent single information (e.g., the mean clinical's change), structured-data can always take into account a set of additional information or sources (e.g., baseline, overall clinical's change, change's direction) at the same time. A first example is the well-known *longitudinal data* problem where phenomena like daily fluctuations of stock prices or quotes, temperatures, chemical and physical measures, growth rates, and clinical courses, are usually represented by means of time-related measurement series. A very simple, popular, and effective approach to analyze these type of data is the so-called *response feature analysis* (Frison and Pocock, 1992; Everitt, 1995; Arndt et al., 2000). In this approach interval-valued data can model the

*change scores* and *total increase scores* associated with longitudinal measures (Senn, Stevens, and Chaturvedi, 2000; Arndt et al., 2000). Intuitively, the change score can be represented by an interval $d_i = [y_T - y_0, y_0 + y_T]$ where $y_{i0}$ indicates the baseline of the treatment response, $y_T$ the last treatment measure, whereas $y_T - y_0$ denotes the overall patient's change. Clearly, the case $y_T - y_0 = 0$ indicates that no changes occurred as an effect of the treatment. A second important application concerns the representation of *uncertainty* and *imprecision* in empirical measurements (Dai, Wang, and Mi, 2013; Ferson et al., 2004). In some circumstances, accuracies and precisions can vary as a function of specific properties of the measurands (Kirkup and Frenkel, 2006) with the consequence that given two measurands, $x$ and $y$, their accuracies and precisions may vary considerably. Clearly, in this situation, the average representations may not well represent all the critical information stored in the data such as, for example, their natural fluctuations. On the contrary, the adoption of an interval-valued approach can provide a valid alternative to represent all the relevant information carried by the measurands. For instance, the two measurands may be represented by two intervals $\tilde{x} = [\bar{x} - \sigma_x, \bar{x} + \sigma_x]$ and $\tilde{y} = [\bar{y} - \sigma_y, \bar{y} + \sigma_x]$ where $\sigma_x$ and $\sigma_y$ denote the uncertainties associated with $x$ and $y$. Of course, interval-valued data can also be applied in many psychological research contexts. For example, in organizational research, studies are often conducted using the so-called within-person approach where information regarding affects, behaviors, interpersonal interactions, work events, and other workplace phenomena are collected over the time (Fisher and To, 2012). Daily diary methods, like the interval-contingent protocol, are the most adopted techniques to regularly collect data related to immediate or recent experiences from the same sample of people for a given interval of time. As a consequence, daily diary measurements can be naturally represented as closed and bounded intervals (Taris, Lange, and Kompier, 2010). This approach is also widely adopted in the assessment of the individuals' observable behaviors where interval data can summarize the coded behaviors (Hartmann, Barrios, and Wood, 2004). Relatedly, interval-valued

data can arise when individual's measurements are collected by means of mouse-tracking instruments that dynamically measure some relevant features of the cognitive processes which are associated with individuals' responses (e.g., Calcagnì and Lombardi, 2014; Johnson et al., 2012). Moreover, interval representations can be also encountered when data are binning or grouped together, for instance, to reduce signal-noise artifacts and/or to improve the quality of the recorded data. These techniques are widespread in cognitive neuroscience studies (Endres et al., 2010). Finally, interval-valued data may arise when no precise single-valued information is available (e.g., in cases of privacy or security concerns, confidentiality requirements, census) and data are naturally incomplete, censored, and masked (Little, 1993). This problem is well-known in health psychology where researchers usually work with partial information (Ogden, 2007). In these cases, interval representations may help in facilitating the recovering of such confidential and sensitive missing data (Kao and Liu, 2000).

### 6.2.2 Formal definitions

The interval $\tilde{a} = [u, v]$ is the set of real numbers $\{x \in \mathbb{R} |\ u \leq x \leq v\}$ where $u$ and $v$ denotes the left and right endpoints of the interval. Two intervals $\tilde{a}$ and $\tilde{b}$ are *equal* if their corresponding endpoints are the same. The interval $\tilde{a}$ is said to be a *degenerated* interval if $u = v$ and in this case the interval simply reduces to the singleton $\tilde{a} = \{u\}$. The *width* of $\tilde{a}$ is defined as $wdt(\tilde{a}) = v - u$, whereas its *midpoint* is $mid(\tilde{a}) = (u + v)/2$. The half-width of $\tilde{a}$ is called the *spread* (or *radius*) of $\tilde{a}$ and is defined as $spr(\tilde{a}) = (v - u)/2$ . Interval-valued data can be easily extended to the multidimensional case. In particular, a $n \times k$ *interval matrix* $\tilde{\mathbf{A}}$ is a matrix whose elements are interval numbers, namely $\tilde{\mathbf{A}} = (\tilde{a}_{ij}) = ([u, v]_{ij})$ with $i = 1...n$ and $j = 1...k$. From a geometrical point of view, the $i$-th row of $\tilde{\mathbf{A}}$ can be represented as a $k$-dimensional hyper-rectangle. More precisely, for $k = 1$, $\tilde{a}$ is a simple interval in $\mathbb{R}$, for $k = 2$, $\tilde{\mathbf{a}}_{1\times 2}$ is a rectangle in $\mathbb{R}^2$, whereas for $k \geq 2$, $\tilde{\mathbf{a}}_{1\times k}$ is an hyper-rectangle in $\mathbb{R}^k$ (see Figure 6.1). The width of $\tilde{\mathbf{A}}$ is the non-negative matrix of widths computed on its interval elements

$\mathbf{A}^w = wdt(\tilde{a}_{ij})$ whereas the spreads of $\tilde{\mathbf{A}}$ is the non-negative matrix of spreads of its interval elements $\mathbf{A}^r = wdt(\tilde{a}_{ij})/2$. Likewise, the midpoint of $\tilde{\mathbf{A}}$ is the matrix containing the midpoints of its interval element $\mathbf{A}^c = mid(\tilde{a}_{ij})$. Further details about the formal properties and operations for interval-valued data can be found in (Moore, 1966).

### 6.2.3 Centre-range parametrization

There are several parametric representations (e.g., centered based, min-max, centre-range) that can be adopted to describe interval-type data (Lima Neto and Carvalho, 2008). Among these, the *centre-range parametrization* allows to describe an interval $\tilde{a}$ by means of its midpoint and spread: $\tilde{a} = (c, r)_{CR}$ where $c = mid(\tilde{a})$ and $r = spr(\tilde{a})$. Unlike other parametric representations for interval-valued data, the CR-parametrization shows some nice features (Blanco-Fernández, Colubi, and García-Bárzana, 2013). Firstly, from a computational perspective, it always ensures well-defined intervals by simply satisfying the non-negative condition $r > 0$.
Secondly, the CR-representation directly works with the parameter space $\mathcal{O}_{\tilde{a}} = \{(c, r) \in \mathbb{R} \times \mathbb{R}^+\}$ of $\tilde{a}$. This would allow to extend many classical statistical approaches to interval-valued data without considering other sophisticated manipulation methods (e.g., interval algebra). Moreover, in the case of multidimensional interval data, the CR-representation can decompose the $n \times m$ *interval matrix* $\tilde{\mathbf{A}}$ into two $n \times m$ *single-valued matrices*, $\mathbf{A^c}$ and $\mathbf{A^r}$, which contain all the parameters involved in $\mathcal{O}_{\tilde{a}}$. In this way, multidimensional interval-valued data may be further manipulated according to classical statistical techniques. Finally, the CR-representation may be very useful especially when intervals are used to describe empirical objects in terms of measurement precisions (by means of $c$) and measurement uncertainty (by means of $r$).

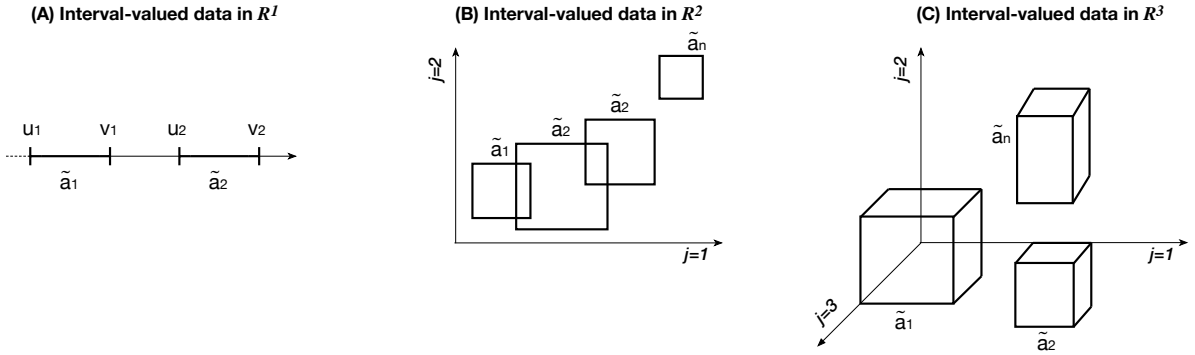



FIGURE 6.1: Graphical representations for interval-valued data: (a) two intervals in $\mathbb{R}^1$, (b) a collection of interval in $\mathbb{R}^2$, (c) intervals in $\mathbb{R}^3$. Note that in the panel (b) a linear relation among the intervals is depicted.

## 6.3 Interval mediation analysis

In this section we illustrate the new multiple mediation model for interval-valued data (IMedA) which is in line with the general framework of symbolic data analysis (e.g., Lima Neto and Carvalho, 2008; Palumbo and Lauro, 2003; Giordani and Kiers, 2004b; Irpino, 2006).

### 6.3.1 General context and motivation

IMedA is a *parallel multiple mediator model* with $k$ mediators where the independent variable affects each mediator which is, in turn, causally linked to the dependent variable but the mediators are assumed not to affect each other. In accordance with the general regression framework in mediation analysis, IMedA uses two regression systems to represent the entire structure of the model. The first system represents the paths between the independent variable and the mediator variables. The second system represents the paths between the mediator variables and the dependent variable. More specifically, from a Least Squares (LS) perspective (Edwards and Lambert, 2007), both the regression systems are first separately estimated and the mediation effects are then computed after the estimated regression coefficients are obtained. Such effects are computed using a

specific procedure for the decomposition of effects in the case of interval-valued data (see next section). Although some empirical contexts may require models involving more than one independent (resp. dependent) variable (usually analyzed via SEMs), in this contribution we preferred to introduce interval mediation analysis for the most simple and widely adopted case first (MacKinnon, 2008; Taylor, MacKinnon, and Tein, 2008).

### 6.3.2 The IMedA model

Let $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ be $n$ (units) $\times$ 1 interval vectors representing the independent and dependent variables, respectively. Let $\tilde{\mathbf{M}}$ be a $n$ (units)$\times$ $k$ (mediators) interval matrix containing the set of mediators.
By adopting the CR parametrization for interval data, the elements of $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}$, and $\tilde{\mathbf{M}}$ can be represented by a collection of $n$ (units) $\times$ 1 single-valued vectors $\mathbf{x^c}$, $\mathbf{x^r}$, $\mathbf{y^c}$, $\mathbf{y^r}$, and $n$ (units) $\times$ $k$ single-valued matrices $\mathbf{M^c}$, $\mathbf{M^r}$. The mediation model for interval-valued data can be expressed by two regression systems as follows:

$$\mathcal{S}_1 : \begin{cases} \mathbf{M^c} = \mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi} + \mathbf{E}^c \\ \mathbf{M^r} = \mathbf{1}\mathbf{A}^r + (\mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi})\boldsymbol{\Pi} + \mathbf{E^r} \end{cases} \tag{6.1}$$

$$\mathcal{S}_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma^r} + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + (\mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma^r})\delta + \boldsymbol{\epsilon}^r \end{cases}$$

For $\mathcal{S}_1$, the matrices $\mathbf{A}^c$, $\mathbf{A}^r$ and $\boldsymbol{\Pi}$ denote $k \times k$ diagonal matrices of intercept terms and coefficients of the ranges, $\mathbf{X}$ is a $n \times 2$ column-wise stacked matrix containing the vectors $\mathbf{x^c}$ and $\mathbf{x^r}$, whereas $\boldsymbol{\Xi}$ is a $2 \times k$ matrix of regression coefficients between the matrix of mediators $\mathbf{M^c}$ and the independent variables $\mathbf{X}$. Finally, $\mathbf{E^c}$ and $\mathbf{E^r}$ are $n \times k$ matrices of residual terms. Similarly, for the second system $\mathcal{S}_2$, the scalars $\alpha^c$, $\alpha^r$, $\delta$ represent the intercept terms and the range coefficient of the model. Moreover, $\boldsymbol{\beta}$ is a $2 \times 1$ vector of regression coefficients quantifying the relation between the independent variables $\mathbf{X}$ and theote that dependent variable $\mathbf{y^c}$ whereas

$\boldsymbol{\gamma}^{\mathbf{c}}$ and $\boldsymbol{\gamma}^{\boldsymbol{r}}$ are $k \times 1$ vectors of regression coefficients between the matrices of mediators $\mathbf{M^c}$ and $\mathbf{M^r}$ and the dependent variable $\mathbf{y^c}$, with $\boldsymbol{\epsilon}^c$ and $\boldsymbol{\epsilon}^r$ being $n \times 1$ vectors of residual terms. Finally, $\mathbf{1}$ denotes matrices (or vectors) of appropriate orders of all ones. Figure 6.2-A (resp. 6.2-B) shows the compact (resp. exploded) conceptual diagram for the IMedA model.

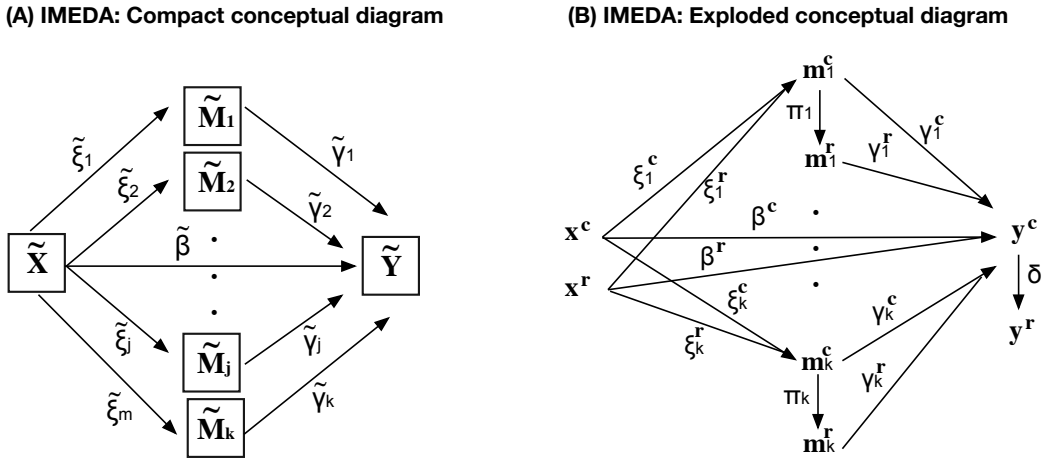



FIGURE 6.2: Conceptual (A) and exploded (B) diagrams for the interval mediation model. Note that the tilde symbol denotes interval variables or coefficients.

## 6.4 Data analysis

### 6.4.1 Parameters estimation

In the IMedA model, the parameters estimates are obtained according to a standard Least Squares (LS) procedure which minimizes the following dissimilarity measures:

$$\begin{aligned} \mathcal{D}_1^2 &= \|\mathbf{M^c} - \mathbf{M^{c*}}\|^2 + \|\mathbf{M^r} - \mathbf{M^{r*}}\|^2 \\ \mathcal{D}_2^2 &= \|\mathbf{y^c} - \mathbf{y^{c*}}\|^2 + \|\mathbf{y^r} - \mathbf{y^{r*}}\|^2 \end{aligned} \tag{6.2}$$

where: $\mathbf{M^{c*}} = \mathbf{1A}^c + \mathbf{X\Xi}$, $\mathbf{M^{r*}} = \mathbf{1A}^r + \mathbf{M^{c*}\Pi}$, $\mathbf{y^{c*}} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\beta} + \mathbf{M^r}\boldsymbol{\gamma}^{\boldsymbol{r}}$, and $\mathbf{y^{r*}} = \mathbf{1}\alpha^r + \mathbf{y^{c*}}\delta$, respectively. To estimate the parameters contained in $\mathbf{A}^c$, $\mathbf{A}^r$, $\mathbf{\Pi}$, $\mathbf{\Xi}$, $\alpha^c$, $\alpha^r$, $\delta$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$, we used the Alternating

Least Squares (ALS) algorithm (Kiers, 2002). By convention, this alternating gradient-descendent algorithm converges when $\|\theta - \hat{\theta}\|^2 \leq \epsilon$ with $\theta$ being the array containing the model's parameters, $\hat{\theta}$ the corresponding estimated array, and $\epsilon$ a small positive quantity, respectively. The detailed iterative ALS solutions of the model are reported in Appendix A whereas the basic estimation algorithm is available as supplementary material to this chapter.

### 6.4.2 Goodness-of-fit indices

The goodness of fit of the IMedA model can be evaluated by considering the following two normalized indices:

$$R^2_M = 1 - \frac{\|\mathbf{M^c} - \mathbf{M^{c*}}\|^2 + \|\mathbf{M^r} - \mathbf{M^{r*}}\|^2}{\|\mathbf{M^c} - \mathbf{1}\,\mathrm{diag}(\overline{\mathbf{M^c}})\|^2 + \|\mathbf{M^r} - \mathbf{1}\,\mathrm{diag}(\overline{\mathbf{M^r}})\|^2}$$
$$R^2_Y = 1 - \frac{\|\mathbf{y^c} - \mathbf{y^{c*}}\|^2 + \|\mathbf{y^r} - \mathbf{y^{r*}}\|^2}{\|\mathbf{y^c} - \overline{\mathbf{y^c}}\|^2 + \|\mathbf{y^r} - \overline{\mathbf{y^r}}\|^2} \tag{6.3}$$

where $\mathrm{diag}(\overline{\mathbf{M^c}})$ and $\mathrm{diag}(\overline{\mathbf{M^r}})$ denote $k \times k$ diagonal matrices containing the column means of the matrices $\mathbf{M^c}$ and $\mathbf{M^r}$, $\mathbf{1}$ is a $n \times k$ matrix of all ones, whereas $\overline{\mathbf{y^c}}$ and $\overline{\mathbf{y^r}}$ denote $n \times 1$ vectors containing the mean values of $\mathbf{y^c}$ and $\mathbf{y^r}$, respectively. Note that $R^2_M$ and $R^2_Y$ take values in $[0, 1]$ and compare the residual sum of squares with the observed total sum of squares. The interpretations of the goodness-of-fit indices are in line with the standard $R^2$ measure adopted in the regression framework.

## 6.5 Analysis of effects

### 6.5.1 Decomposition of effects

The decomposition of effects allows to quantify the amount of effect produced by the mediators on the relation between the dependent variable and the independent variable (indirect effect) and the residual total effect

between the two variables when the mediators are held constant (direct effect) (Alwin and Hauser, 1975). In particular, considering the regression systems $\mathcal{S}_1$ and $\mathcal{S}_2$ in Eq. 6.1, the *total effect* (TE) of $\tilde{\mathbf{x}}$ on $\tilde{\mathbf{y}}$ can be partitioned into *direct* (DE) and *indirect* (IE) effects. More precisely, the indirect effect (IE) of $\tilde{\mathbf{x}}$ on $\tilde{\mathbf{y}}$ is the amount that $\tilde{\mathbf{y}}$ is expected to change as $\tilde{\mathbf{x}}$ changes as a result of the effect of $\tilde{\mathbf{x}}$ on $\tilde{\mathbf{M}}$ which, in turn, affects $\tilde{\mathbf{y}}$. On the contrary, the direct effect (DE) is usually regarded as that part of the total effect (TE) which is not transmitted through $\tilde{\mathbf{M}}$ and quantifies how much a change in $\tilde{\mathbf{x}}$ affects $\tilde{\mathbf{y}}$ independent of its effect on $\tilde{\mathbf{M}}$ (Alwin and Hauser, 1975; Stolzenberg, 1980; Sobel, 1990). As a consequence, the total effect (TE) can be expressed as TE = DE + IE which represents the basic decomposition rule for TE. In the IMedA model (see Figure 6.1-B), this basic rule can be generalized as shown in Table 6.1, where $\text{TE}_c$, $\text{TE}_r$, $\text{DE}_c$, $\text{DE}_r$, $\text{IE}_c$ and $\text{IE}_r$ represent the interval components of TE, DE and IE for centers and ranges, respectively. The terms $\text{IE}^{()}_{c/c}$, $\text{IE}^{()}_{r/r}$, $\text{IE}^{()}_{c/r}$ and $\text{IE}^{()}_{r/c}$ indicate the interval *elementary indirect effects* of the model.

| |
|---|
| $\text{TE} = \text{DE} + \text{IE}$ |
| $\text{TE}_c + \text{TE}_r = [\text{DE}_c + \text{DE}_r] + [\text{IE}_c + \text{IE}_r]$ |
| $\text{TE}_c + \text{TE}_r = [(\text{DE}_c^{y^c} + \text{DE}_c^{y^r}) + (\text{DE}_r^{y^c} + \text{DE}_r^{y^r})] + [(\text{IE}_c^{y^c} + \text{IE}_c^{y^r}) + (\text{IE}_r^{y^c} + \text{IE}_r^{y^r})]$ |
| with: |
| $\text{IE}_c^{y^c} + \text{IE}_c^{y^r} = (\text{IE}_{c/c}^{y^c} + \text{IE}_{c/r}^{y^c}) + (\text{IE}_{c/c}^{y^r} + \text{IE}_{c/r}^{y^r})$ |
| $\text{IE}_r^{y^c} + \text{IE}_r^{y^r} = (\text{IE}_{r/c}^{y^c} + \text{IE}_{r/c}^{y^c}) + (\text{IE}_{r/r}^{y^r} + \text{IE}_{r/r}^{y^r})$ |

TABLE 6.1: Decomposition rules for the IMedA model involved by the CR-parametrization

It is straightforward to note that in IMedA both the centers and ranges contribute in the total effect of the model. However, because of the characteristics of the IMedA model, we can only use the regression equation of $\mathbf{y^c}$ to derive all the effects of the model, whereas the effects related to the regression equation of $\mathbf{y^r}$ (i.e., $\text{DE}_c^{y^r}$, $\text{DE}_r^{y^r}$, $\text{IE}_c^{y^r}$, and $\text{IE}_r^{y^r}$) are subsequently computed via $\delta$. As a result of this property, the third line of the decomposition rules in Table 6.1 can be re-written as:

$$\text{TE}_c + \text{TE}_r = (\text{DE}_c^{y^c} + \text{DE}_r^{y^c} + \text{IE}_c^{y^c} + \text{IE}_r^{y^c}) + \delta(\text{DE}_c^{y^c} + \text{DE}_r^{y^c} + \text{IE}_c^{y^c} + \text{IE}_r^{y^c}) \quad (6.4)$$

where $\delta(\text{DE}_c^{y^c} + \text{DE}_r^{y^c} + \text{IE}_c^{y^c} + \text{IE}_r^{y^c}) = \text{DE}_c^{y^r} + \text{DE}_r^{y^r} + \text{IE}_c^{y^r} + \text{IE}_r^{y^r}$. Likewise,

the elementary indirect effects can also be re-written considering the coefficient $\delta$. By applying the decomposition procedure for interval mediation analysis (see Appendix B), we can formally write all the terms shown in Eq. 6.4, as follows:

$$\begin{aligned}
&\mathrm{DE}_c^{y^c} = \hat{\beta}^c \qquad \mathrm{DE}_r^{y^c} = \hat{\beta}^r \\
&\mathrm{IE}_c^{y^c} = \mathrm{IE}_{c/c}^{y^c} + \mathrm{IE}_{c/r}^{y^c} = (\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{cT})\mathbf{1}_k + (\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{rT} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k \\
&\mathrm{IE}_r^{y^c} = \mathrm{IE}_{r/c}^{y^c} + \mathrm{IE}_{r/r}^{y^c} = (\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{cT})\mathbf{1}_k + (\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{rT} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k
\end{aligned} \tag{6.5}$$

where $\hat{\beta}^c$ and $\hat{\beta}^r$ are the estimated parameters contained in $\widehat{\boldsymbol{\beta}}$; $\hat{\boldsymbol{\xi}}^c$ and $\hat{\boldsymbol{\xi}}^r$ are $1 \times k$ row-vectors of the estimated matrix $\widehat{\boldsymbol{\Xi}}$; $\widehat{\boldsymbol{\gamma}}^c$, $\widehat{\boldsymbol{\gamma}}^r$, and $\widehat{\boldsymbol{\Pi}}$ are the estimated parameters previously defined; $\mathbf{1}_k$ is a $k \times 1$ vector of all ones whereas $\circ$ denotes the usual Hadamard product. Note that, the term $\mathbf{1}_k$ in Eq. 6.5 allows to compute the *total indirect effect*, that is to say, the sum of the *elementary indirect effects*. These are the indirect effects that are associated with each of the $k$ mediators in the model. In particular, if we omit the term $\mathbf{1}_k$ from the equations of the effects, we obtain the specific indirect effects separately for each mediator in the model. To summarize, the complete effects involved in the interval mediation analysis are reported in table 6.2.

| | | |
|---|---|---|
| $\mathrm{DE}_c$ | = | $\hat{\beta}^c + \hat{\delta}\hat{\beta}^c$ |
| $\mathrm{DE}_r$ | = | $\hat{\beta}^r + \hat{\delta}\hat{\beta}^r$ |
| $\mathrm{IE}_c$ | = | $[(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{cT})\mathbf{1}_k + (\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{rT} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k] + \hat{\delta}[(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{cT})\mathbf{1}_k + (\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{rT} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k]$ |
| $\mathrm{IE}_r$ | = | $[(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{cT})\mathbf{1}_k + (\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{rT} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k] + \hat{\delta}[(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{cT})\mathbf{1}_k + (\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{rT} \circ \hat{\boldsymbol{\Pi}}^T)\mathbf{1}_k]$ |

TABLE 6.2: Decomposition of effects in the IMedA model

### 6.5.2 Evaluating the size of the effects

In this section we introduce some descriptive indices to evaluate the size of the decomposed effects of the IMedA model. In the context of mediation analysis, several different indices have been defined to quantify the size of the decomposed effects, such as ratio measures (e.g., the proportion mediated index), standardized regression parameters, and indices based on the variance decomposition (e.g., partial $r^2$, residual-based indices. See:

Preacher and Kelley, 2011; Fairchild et al., 2009). Because of its simplicity, variance decomposition becomes the most popular criterion in defining mediation effect size (although a number of shortcomings have been highlighted elsewhere, see MacKinnon, 2008). In this contribution, we resort to a set of particular indices based on the decomposition of the explained variance. Unlike other $R^2$-based decompositions which make use of the so-called commonality analysis (Seibold and McPhee, 1979), our indices are defined considering the so-called *reduced system* which is obtained by merging $\mathcal{S}_1$ with $\mathcal{S}_2$ (see Appendix C). In this new equations system, the dependent variable $\tilde{\mathbf{y}}$ is modeled as a function of all the pathways expressed by the IMedA model and its variance can be partitioned according to a dedicated regression based decomposition procedure (Mood and Graybill, 1974; Fields, 2003). This would allow to show the contribution of the direct and indirect effects in modeling the variance of $\tilde{\mathbf{y}}$, as follows:

$$\mathrm{DE}_{\sigma^2_{\tilde{\mathbf{y}}}} = \begin{bmatrix} \mathrm{cov}(\mathbf{y^c}, \mathbf{x^c}\hat{\beta}^c) + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^r}\hat{\beta}^r) + \\ \mathrm{cov}(\mathbf{y^r}, \hat{\delta}\mathbf{x^c}\hat{\beta}^c) + \mathrm{cov}(\mathbf{y^r}, \hat{\delta}\mathbf{x^r}\hat{\beta}^r) \end{bmatrix} \cdot \begin{bmatrix} \mathrm{var}(\mathbf{y^c}) + \\ \mathrm{var}(\mathbf{y^r}) \end{bmatrix}^{-1} \tag{6.6}$$

$$\mathrm{IE}_{\sigma^2_{\tilde{\mathbf{y}}}} = \begin{bmatrix} \mathrm{cov}(\mathbf{y^c}, \mathbf{x^c}(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{c\,T})\mathbf{1}_k) + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^c}(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{r\,T} \circ \hat{\mathbf{\Pi}}^T)\mathbf{1}_k) + \\ \mathrm{cov}(\mathbf{y^c}, \mathbf{x^r}(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{c\,T})\mathbf{1}_k) + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^r}(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{r\,T} \circ \hat{\mathbf{\Pi}}^T)\mathbf{1}_k) + \\ \mathrm{cov}(\mathbf{y^r}, \hat{\delta}\mathbf{x^c}(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{c\,T})\mathbf{1}_k) + \mathrm{cov}(\mathbf{y^r}, \hat{\delta}\mathbf{x^c}(\hat{\boldsymbol{\xi}}^c \circ \hat{\boldsymbol{\gamma}}^{r\,T} \circ \hat{\mathbf{\Pi}}^T)\mathbf{1}_k) + \\ \mathrm{cov}(\mathbf{y^r}, \hat{\delta}\mathbf{x^r}(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{c\,T})\mathbf{1}_k) + \mathrm{cov}(\mathbf{y^r}, \hat{\delta}\mathbf{x^r}(\hat{\boldsymbol{\xi}}^r \circ \hat{\boldsymbol{\gamma}}^{r\,T} \circ \hat{\mathbf{\Pi}}^T)\mathbf{1}_k) \end{bmatrix} \cdot \begin{bmatrix} \mathrm{var}(\mathbf{y^c}) + \\ \mathrm{var}(\mathbf{y^r}) \end{bmatrix}^{-1} \tag{6.7}$$

$$\mathrm{RES}_{\sigma^2_{\tilde{\mathbf{y}}}} = \begin{bmatrix} \mathrm{cov}(\mathbf{y^c}, \hat{\mathbf{E}}^c\hat{\gamma}^c) + \mathrm{cov}(\mathbf{y^c}, \hat{\mathbf{E}}^r\hat{\gamma}^r) + \\ \mathrm{cov}(\mathbf{y^r}, \hat{\mathbf{E}}^c\hat{\gamma}^c\hat{\delta}) + \mathrm{cov}(\mathbf{y^r}, \hat{\mathbf{E}}^r\hat{\gamma}^r\hat{\delta}) \end{bmatrix} \cdot \begin{bmatrix} \mathrm{var}(\mathbf{y^c}) + \\ \mathrm{var}(\mathbf{y^r}) \end{bmatrix}^{-1} \tag{6.8}$$

Note that the quantity $\mathrm{DE}_{\sigma^2_{\tilde{\mathbf{y}}}}$ contains the components of the variance referred to the effects $\mathrm{DE}_c$ and $\mathrm{DE}_r$ for both $\mathbf{y^c}$ and $\mathbf{y^r}$. $\mathrm{IE}_{\sigma^2_{\tilde{\mathbf{y}}}}$ contains the indirect effects $\mathrm{IE}_c$ and $\mathrm{IE}_r$ together with the partial indirect effects for the $\mathbf{y^c}$ and $\mathbf{y^r}$ components of the model, whereas $\mathrm{RES}_{\sigma^2_{\tilde{\mathbf{y}}}}$ refers to the residual terms of the models. Finally, because the sum $\Lambda = \mathrm{DE}_{\sigma^2_{\tilde{\mathbf{y}}}} + \mathrm{IE}_{\sigma^2_{\tilde{\mathbf{y}}}} + \mathrm{RES}_{\sigma^2_{\tilde{\mathbf{y}}}}$ is equal to the observed variability explained by all the pathways in the IMedA model, we can define the following two indices:

$$\lambda_{\mathrm{DE}} = |\mathrm{DE}_{\sigma^2_{\tilde{\mathbf{y}}}}| \cdot \Lambda^{-1} \qquad \lambda_{\mathrm{IE}} = |\mathrm{IE}_{\sigma^2_{\tilde{\mathbf{y}}}}| \cdot \Lambda^{-1} \quad \text{with} \quad \lambda_{\mathrm{DE}}, \lambda_{\mathrm{IE}} \in [0, 1] \tag{6.9}$$

where $|.|$ indicates the absolute value. In our context, $\lambda_{\text{DE}}$ and $\lambda_{\text{IE}}$ can be seen as the proportion of the observed variability explained by the effects of the model which is exclusively due to either the direct effect ($\lambda_{\text{DE}}$) or the indirect effect ($\lambda_{\text{IE}}$). Note that when $\lambda_{\text{DE}}$ approaches $0$ (and consequently, $\lambda_{\text{IE}}$ approaches 1) the representation reduces to the so-called *full mediation* case in which the variance explained by the effects is exclusively due to the mediators in the model. By contrast, if $\lambda_{\text{DE}}$ approaches $1$ (consequently, $\lambda_{\text{IE}}$ approaches $0$), then the mediation model is said to be *ill-posed* because mediators do not contribute in explaining the observed variability in the model. Moreover, by omitting the term $\mathbf{1}_k$ in the Eq. 6.7, we obtain $k$ partial indices $\lambda^1_{\text{IE}}, ..., \lambda^j_{\text{IE}}, ..., \lambda^k_{\text{IE}}$ (with $\sum_j^k \lambda^j_{\text{IE}} = \lambda_{\text{IE}}$) where each index $\lambda^j_{\text{IE}}$ represents the proportion of the variance explained by that effect which is specifically due to the corresponding mediator $\tilde{\mathbf{M}}_j$.

## 6.6 Simulation study

The aim of this simulation study is twofold. First, we will evaluate the properties of the estimators of the proposed IMedA-ALS algorithm. The least squares estimators for mediation analysis and symbolic regression have been extensively studied in several simulation works (MacKinnon et al., 2002; Preacher and Hayes, 2008; Zhang and Wang, 2013; Nkurunziza and Ejaz Ahmed, 2011; Lima Neto and Carvalho, 2008). Overall, these results confirmed that OLS estimators for mediation models are correct, efficient, and accurate. However, although our model inherits all the classical least squares properties, in the present study we preferred to evaluate the performances of the IMedA-ALS algorithm for the sake of completeness and to further provide converging results.

Second, we will re-analyse the same simulated data by means of two alternative methods, namely a standard SEM approach and a regression-based mediation approach for mediation analyses. Because such methodologies have been widely studied in the mediation literature (Little, 1993; MacKinnon, 2008), one could ask whether they can be also adopted in the case

of interval-valued data with minimal changes in their model's representations. Therefore, the second aim of our simulation study is to understand whether these well-known approaches can appropriately reproduce the interval-valued pathways generated by the original IMedA model representation (see Figure 6.2-B).
The simulation study is conducted for the case of a one mediator ($m = 1$) and two mediators ($m = 2$) models, respectively. However, because the results for the $m = 2$ case largely mirrored those of the simpler $m = 1$ case, in the following paragraph we will discuss the latter case only.

### 6.6.1 Design and procedure

Two factors were systematically varied in a complete two-factorial design:

(i) the sample size ($n$) at four levels: 50, 250, 500, 1000;

(ii) the amount of noise ($e$) at four levels: 0.10, 0.30, 0.50, 0.70. Factor $e$ is defined as the proportion of the total variance in the data that is not accounted by the IMedA model. Technically, the proportion of error in the data is computed by modeling the variances of the error terms in the IMedA model using a predefined set of values stored in two matrices, $\mathbf{H}^E$ and $\mathbf{H}^\epsilon$, of error variances associated to the IMedA $\mathcal{S}_1$ and $\mathcal{S}_2$ regression systems, respectively. These values were defined according to a previous simulation study and reflect the condition that the proportions of explained variance accounted by the IMedA model always equal to $1 - e_k$ ($k = 1 \dots 4$).

Now we are in the position to provide all the details of our simulation design. Let $n_k$ and $e_k$ be distinct levels of the factors $n$ and $e$ respectively. The following procedural steps were repeated 1000 times ($Q = 1000$) for each of the 16 combinations of levels of the simulation design:

(a) Generate the interval data matrix $\tilde{\mathbf{x}}_{n_k \times 2} = ([u, v]_{ij})$ from the uniform distribution $\mathcal{U}(1, 10)$ with $[u < v]_{ij}$. Next, obtain $\mathbf{x^c}$ and $\mathbf{x^r}$ via the CR-parametrization on $\tilde{\mathbf{x}}_{n_k \times 2}$;

(b) Generate the mediator variables $\mathbf{M^c}_{(n_k \times m)}$ and $\mathbf{M^r}_{(n_k \times m)}$ (with $m = 1$) by applying the regression system $\mathcal{S}_1$ with $\mathbf{E^c} \sim N(0, \mathbf{H}^E_{k,c})$ and $\mathbf{E^r} \sim N(0, \mathbf{H}^E_{k,r})$ with the following parameters: $\mathbf{A^c} = 4.8$, $\mathbf{A^r} = 3.1$, $\boldsymbol{\Xi} = (2.7, 4.1)^T$, $\boldsymbol{\Pi} = 2.04$;

(c) Estimate the parameters $\hat{\mathbf{A}}^{\mathbf{c}}_q$, $\hat{\mathbf{A}}^{\mathbf{r}}_q$, $\widehat{\boldsymbol{\Xi}}_q$, $\widehat{\boldsymbol{\Pi}}_q$ of the system $\mathcal{S}_1$ for the $q$-th sample by means of the IMedA-ALS estimators (see Appendix A);

(d) Generate the dependent variables $\mathbf{y^c}_{(n_k \times 1)}$ and $\mathbf{y^r}_{(n_k \times 1)}$ by applying the regression system $\mathcal{S}_2$ with $\boldsymbol{\epsilon^c} \sim N(0, \mathbf{H}^{\epsilon}_{k,c})$, $\boldsymbol{\epsilon^r} \sim N(0, \mathbf{H}^{\epsilon}_{k,r})$ and the following parameters: $\alpha^{\mathbf{c}} = 3.0$, $\alpha^{\mathbf{r}} = -5.3$, $\boldsymbol{\beta} = (2.3, 1.9)^T$, $\gamma^{\mathbf{c}} = 1.9$, $\gamma^{\mathbf{r}} = 0.9$, and $\delta = -3.25$.

(e) Estimate the parameters $\hat{\alpha}^{\mathbf{c}}_q$, $\hat{\alpha}^{\mathbf{r}}_q$, $\hat{\boldsymbol{\beta}}_q$, $\hat{\gamma}^{\mathbf{c}}_q$, $\hat{\gamma}^{\mathbf{r}}_q$, and $\hat{\delta}_q$ of the system $\mathcal{S}_2$ for the $q$-th sample by means of the IMedA-ALS estimators (see Appendix A);

(f) Save the estimates and proceed until $q = 1000$.

This procedure was used to generate the estimated distributions of the regression parameters for each combination of levels of the simulation design. The whole procedure generated a total of $1000 \times 4 \times 4 = 16000$ new data matrices as well as an equivalent number of parameters.

### 6.6.2 Outcome measures

The sample results were evaluated considering the following *global measures* that give information about the overall performance of the IMedA-ALS estimates:

(a) *average root mean square error* (AMSE) computed as:

$$AMSE = Q^{-1} \sum_q \sqrt{J^{-1} \sum_j \left[ (\widehat{\theta}_{qj} - \theta_{qj}) \cdot {\theta_{qj}}^{-1} \right]^2}$$

with $\theta_q$ and $\widehat{\theta}_q$ being the arrays of parameters of the true and estimated model, respectively. Low values of AMSE indicate that the estimators accurately reproduce the true parameter values;

(b) *proportion of agreement* (PA) index computed as:

$$PA = Q^{-1} \sum_{q} \left[ 1 - \left( \|\widehat{\theta}_q - \theta_q\|^2 \cdot (\|\theta_q\|^2)^{-1} \right) \right] 100$$

The index takes values in $[0, 100]$ and assesses how much the estimated array of parameters $\widehat{\theta}$ resembles the true array $\theta$ (Timmerman and Kiers, 2002). When PA is closed to 100 the estimated array $\widehat{\theta}$ perfectly recovers the true array $\theta$.

### 6.6.3 Results

The first column of Tables 6.3 reports the results of the simulation study. As expected, the AMSE index decreased almost linearly with increasing sample sizes whereas increased with increasing perturbation terms $e$. On the contrary, the PA index increased with increasing sample size. In particular, for $n \geq 250$, PA became unaffected by factor $e$. Overall, the IMedA-ALS algorithm was good and very stable also in cases of high noise terms. Clearly, these results confirmed how the ALS algorithm upon which IMedA is based generally shows highly robust estimates. To summarize, IMedA-ALS always produced excellent estimates when the amount of noise in the data was low ($e = 0.1$) or moderate ($e = 0.3$). Moreover, also in cases of large ($e = 0.5$) or extreme amount of noise ($e = 0.7$), IMedA-ALS still showed undistorted estimates at least when $n > 50$. By contrast, for small sample sizes the performance decreased according to the amount of noise.

### 6.6.4 Further analysis and results

We also evaluated the performances of a standard SEM approach for single-valued data and a least squares procedure for standard mediation analysis in reconstructing the data generated in the previous

| $n$, $e$ | IMedA-ALS | | SEM-ML | | SEM-WLS | | 2SMA | |
|---|---|---|---|---|---|---|---|---|
| | AMSE | PA | AMSE | PA | AMSE | PA | AMSE | PA |
| $n = 50$ | | | | | | | | |
| $e_1$ | 0.11 | 99.45 | 0.44 | 84.36 | 14.67 | 65.89 | 0.44 | 84.37 |
| $e_2$ | 0.21 | 98.00 | 0.52 | 80.58 | 24.38 | 66.95 | 0.52 | 80.76 |
| $e_3$ | 0.28 | 95.13 | 0.53 | 80.00 | 21.24 | 61.93 | 0.53 | 80.25 |
| $e_4$ | 0.31 | 93.55 | 0.53 | 79.71 | 28.77 | 63.82 | 0.52 | 80.29 |
| $n = 250$ | | | | | | | | |
| $e_1$ | 0.05 | 99.88 | 0.19 | 96.65 | 10.74 | 75.04 | 0.19 | 96.68 |
| $e_2$ | 0.08 | 99.64 | 0.22 | 95.79 | 12.26 | 71.98 | 0.21 | 95.99 |
| $e_3$ | 0.11 | 99.44 | 0.25 | 94.95 | 15.53 | 73.49 | 0.23 | 95.34 |
| $e_4$ | 0.14 | 99.12 | 0.26 | 94.66 | 25.21 | 71.49 | 0.23 | 95.39 |
| $n = 500$ | | | | | | | | |
| $e_1$ | 0.03 | 99.95 | 0.14 | 98.26 | 6.88 | 79.19 | 0.14 | 98.29 |
| $e_2$ | 0.06 | 99.85 | 0.16 | 97.85 | 13.62 | 77.27 | 0.15 | 98.05 |
| $e_3$ | 0.08 | 99.73 | 0.19 | 97.27 | 13.52 | 76.35 | 0.16 | 97.67 |
| $e_4$ | 0.09 | 99.66 | 0.20 | 97.00 | 26.44 | 78.81 | 0.16 | 97.73 |
| $n = 1000$ | | | | | | | | |
| $e_1$ | 0.02 | 99.98 | 0.10 | 99.17 | 6.79 | 86.15 | 0.10 | 99.20 |
| $e_2$ | 0.04 | 99.92 | 0.12 | 98.78 | 14.16 | 83.31 | 0.11 | 98.98 |
| $e_3$ | 0.05 | 99.87 | 0.15 | 98.37 | 6.97 | 82.51 | 0.12 | 98.77 |
| $e_4$ | 0.07 | 99.84 | 0.16 | 98.27 | 26.89 | 83.50 | 0.11 | 99.00 |

TABLE 6.3: Monte Carlo study: Percentage of agreement (PA) index and average root mean square errors (AMSE) for the array of parameters of the single mediation model ($m = 1$)

simulation design (see also Fig. 6.2-B). In particular, SEM model fitting and estimation were implemented through the Lavaan R-package (Rosseel, 2012) using the standard SEM representation for mediation analysis with both ML and WLS estimation procedures. On the contrary, the standard regression approach, named 2SMA, was instead implemented through a combination of Matlab scripts that modeled a step-by-step regression procedure. Finally, the AMSE and PA measures were computed for each condition of the simulation design and each of the three estimation procedures (SEM-ML, SEM-WLS, 2SMA). The second, third, and fourth columns of the Table 6.3 report the ensuing results. An inspection of Table 6.3 suggests how SEM-ML and 2SMA produced comparable results in estimating the IMedA parameters. As expected, both the estimation algorithms produced lower errors with increasing sample size. By contrast, SEM-WLS always showed lower performances and accuracies than the other two methods. However, it should be noted that all performances of SEM-ML,

SEM-WLS, and 2SMA were not as good as those obtained using the original IMedA-ALS algorithm. Indeed, although SEM-ML and 2SMA provided acceptable results in resembling the true model structure provided by the IMedA pathways, they still showed an higher value of AMSE for each condition of the simulation design even when increasing sample sizes were considered. By and large, this can reflect relevant structural differences in the modeling and estimation procedures adopted by the other two alternative approaches.

## 6.7 Empirical examples

By way of illustration we consider two data sets analyzed using the IMedA approach. The first study evaluates how the relation between role and work-related burnout can be mediated by some organizational variables (e.g., job satisfaction, job tenure). The second study describes a mediation model in human computer interaction (HCI) studies. In particular, we analysed whether the relationship between the perceived usability and perceived beauty of the online system ESSE-3 is mediated by hedonic qualities and familiarity.

### 6.7.1 Role and work-related burnout

*General context and motivation*. Role is an important variable in many organizational research settings (Sawyer, 1992; Toderi et al., 2013). By and large, role is considered a relevant dimension in predicting employee health and, more in general, organizational stress (Bliese and Castro, 2000). Similarly, work-related burnout corresponds to a protracted individual response to a set of emotional and interpersonal stressors which are presented in the organization (Ashforth and Mael, 1989). In this illustrative example we tested a model in which the basic linear relation between role (R) and work-related burnout (WB) was evaluated by considering job satisfaction (S) and workload (WL) as mediators.

*Data and measures*. In this first application we used some recently published data (Avanzi et al., 2012). In particular, data refers to Italian teachers who participated in a psychosocial risk assessment evaluation conducted in five schools in the Trentino region (northeast of Italy). Data were collected by means of a specific questionnaire which was administered in two different occasions (T1 and T2). The sample was composed of $n = 140$ teachers (83% females), with a mean age of 41.2 years (ranging from 23 to 62 with SD = 10.7) and a mean workload of 19.2 years (ranging from 2 to 37 with SD = 10.3). Because in this case the variables are collected longitudinally, we decided to pre-process the T1-T2 variables according to the *response feature analysis* (Senn, Stevens, and Chaturvedi, 2000; Everitt, 1995). In particular, considering $\min(T1, T2)$ and $\max(T1, T2)$ as the lower and the upper bounds of intervals, centers and ranges were defined according to the CR-parametrization. Note that in this context, the range of an interval directly refers to a *change score* of the measurements collected over the time.

*Data analysis and results*. The IMedA algorithm required 781 iterations to estimate the parameters in the $\mathcal{S}_1$ system and 1084 iterations to estimate those in the $\mathcal{S}_2$ system. The performance of the IMedA model was moderately good ($R^2_M = 0.39$ and $R^2_Y = 0.48$). Table 6.4 shows the final estimated parameters together with the corresponding direct and indirect effects for the model. A quick inspection of Table 6.4 reveals that the variance explained by all the pathways of the model was $\Lambda = 0.60$ where the direct effect, the first mediator (WL), and the second mediator (S) contributed for $22\%$, $36\%$, and $42\%$ of the variance, respectively. The results indicate that the centers of R did not have a linear impact on WB ($\beta^c = 0.06$) whereas the ranges of R showed a weak but significant impact on WB ($\beta^r = 0.21$). The centers of R were negatively related to WL ($\xi^c_1 = -0.26$) but positively associated with S ($\xi^c_2 = 0.68$). Considering the second regression system $\mathcal{S}_2$, the centers and ranges of WL were positively associated with WB

($\gamma_1^c = 0.68$ and $\gamma_1^r = 0.14$), whereas the centers and ranges of S were negatively related to WB ($\gamma_2^c = -0.38$, $\gamma_2^r = -0.02$). Interestingly, in this example the direct effect of R was mainly produced by the ranges ($\text{DE}_r = 0.19$). By contrast, the centers did not show a significant effect ($\text{DE}_c = 0.06$). The results also suggest that the indirect effects for the centers and ranges were both significant. In particular, the indirect effect through the centers of WL ($\text{IE}_c = -0.16$) was negative and weaker than the corresponding effect through S ($\text{IE}_c = -0.27$). Instead, the indirect effect through the ranges of WL ($\text{IE}_r = 0.12$) and S ($\text{IE}_r = -0.10$) showed opposite directions and both the effects were significant. Overall, considering the centers component, the results indicated that R has a protective impact on WB by reducing the positive relation between WL and WB. In a similar way, R improves the negative relation between S and WB, that is to say, the more the role is perceived as clear by the teacher, the higher is the perceived work satisfaction and the lower is the perceived work-related burnout. By contrast, considering the ranges component, R improves the positive impact of WL on WB whereas it makes the relation between S and WB vanish. Because in this context ranges are interpreted as change scores, this result would possibly highlight how the more the teachers experience unstable clarity of role, the more they experience workload and work-related burnout.

### 6.7.2 Beauty and usability in HCI perspective

*General context and motivation*. The interplay between perceived beauty and usability has been extensively studied in HCI perspective (Hassenzahl, 2004). In this context, beauty has been defined as "a predominantly affect-driven evaluative response to the visual Gestalt of an object" (Hassenzahl, 2008) whereas usability has been identified as the ability of a product to fulfill the user's needs and to support the so-called "do-goals" (Hassenzahl and Monk, 2010). Over the years,

| | | first mediator (WL) | | second mediator (S) | |
|---|---|---|---|---|---|
| | | values | 95% CIs | values | 95% CIs |
| Model parameters (M) | $\mathbf{A}^c$ | 3.92 | (2.88,4.88) | 0.81 | (-0.45,1.94) |
| | $\mathbf{A}^r$ | 0.36 | (-1.59,2.46) | 0.17 | (-0.99,1.42) |
| | $\mathbf{\Pi}$ | -0.08 | (-0.73,0.69) | 0.08 | (-0.27,0.38) |
| | $\boldsymbol{\xi}^c$ | -0.26 | (-0.50,-0.09) | 0.68 | (0.41,0.96) |
| | $\boldsymbol{\xi}^r$ | 0.17 | (0.09,0.55) | 0.18 | (-0.14,0.57) |
| Model parameters (Y) | $\alpha^c$ | 1.50 | (-0.22,2.37) | | |
| | $\alpha^l$ | 0.18 | (0.03,0.38) | | |
| | $\delta$ | 0.03 | (0.05,0.25) | | |
| | $\beta^c$ | 0.06 | (-0.06,0.39) | | |
| | $\beta^r$ | 0.21 | (0.04,0.46) | | |
| | $\boldsymbol{\gamma}^c$ | 0.68 | (0.54,0.81) | -0.38 | (-0.48,-0.24) |
| | $\boldsymbol{\gamma}^r$ | 0.14 | (0.07,0.39) | -0.02 | (-0.16,0.07) |
| Effects | $\text{DE}_c$ | 0.06 | (-0.06,0.38) | | |
| | $\text{DE}_r$ | 0.19 | (0.05,0.46) | | |
| | $\text{IE}_c$ | -0.16 | (-0.33,-0.02) | -0.27 | (-0.40,-0.12) |
| | $\text{IE}_r$ | 0.12 | (0.02,0.39) | -0.10 | (-0.24,-0.01) |

*Note:* $R^2_M = 0.39$, $R^2_Y = 0.48$, $\lambda_{\text{DE}} = 0.22$, $\lambda^1_{\text{IE}} = 0.36$, $\lambda^2_{\text{IE}} = 0.42$, CIs indicate the 95% confidence intervals obtained by bias-corrected and accelerating (BCa) bootstrap with 5000 bootstrap samples.

TABLE 6.4: First case study: Estimated model parameters and effects

a number of experiments have explored the basic question "what is beautiful is usable" with mixed and sometimes opposing results (Hamborg, Hülsmann, and Kaspar, 2014; Tractinsky, Katz, and Ikar, 2000). Among others, the well-known Hassenzahl's model offers an integrated framework in which the relationship between usability, beauty, and other important product attributes can be evaluated. By adopting such model, we analysed the perceived beauty and usability of a software called ESSE-3. In particular, ESSE-3 is a dedicated online system developed by CINECA and KION for supporting the digitalization of Italian universities (CINECA and Kion, 1999). It provides a specific student management system through which universities and institutions can supply academic services such as registration, examinations, e-learning, and portfolio to students, researchers, and professors. Although ESSE-3 is used by almost all Italian universities, a growing debate about its usability and overall quality has been carried out over the recent years. To shed light on this debate, we studied whether

the relationship between usability (PQ) and perceived beauty (B) of the system ESSE-3 is mediated by its hedonic qualities (HQ) and the user's experience with this system (F). Note that the hedonic quality (HQ) is referred as the ability of a product to stimulate user's pleasure in use and ownership (Hassenzahl and Monk, 2010) whereas user's experience evaluates the user's familiarity with the product in use.

*Data and measures*. Fifty individuals (16 men, 34 women) participated in the study. All participants were 1st or 4nd year students at the University of Trento (Italy) that reported continuous experience with the ESSE-3 system. The sample's mean age was 22 years (ranging from 19 to 48) with SD equals to 4.03. All the variables involved in the study were measured by the AttracDiff2 questionnaire with bipolar items measured with 7-point scales (Hassenzahl, Burmester, and Koller, 2003). In particular, usability refers to the "Pragmatic Quality" (PQ) dimension of the questionnaire whereas hedonic quality to the "Hedonic Quality-Stimulation" (HQ) dimension. Note that, beauty (B) and familiarity (F) were obtained by using two single 7-point scales ranging from hugly-beauty and unknown-known, respectively. The self-reported responses were collected by adopting the DYFRAT system (Calcagnì and Lombardi, 2014), a novel graphical interface to dynamically measure some properties of the rating responses. In particular, using the mouse-tracking paradigm, the system collects the streaming x–y coordinates of the computer mouse together with the response time of the final responses (Johnson et al., 2012).[1] Subsequently, for each participant's response, the streaming data was modeled as an histogram whose density expresses the temporal and spatial uncertainty of the final response (see Figure 6.3).

[1]More technically, for each item, participants were told that a bended scale with seven response levels would be presented on the screen, and that they were asked to choose which of these levels was the most appropriate for the presented item. After participants clicked a start button, a window with the scale appeared at the top of the screen while the cursor was allocated to the center of the screen. Participants gave their responses by mouse-clicking the chosen level of the scale and, in the meanwhile, the DYFRAT system recorded the x–y coordinates of the computer mouse with the their time.

Because there exist a structural relationship between interval-valued data and histogram-valued data (Billard and Diday, 2003), we easily derived intervals from the *supports* of the observed histograms. In this way, interval-valued data directly codify spatial uncertainties that are related to the mouse-movements underlying the final individual's response.

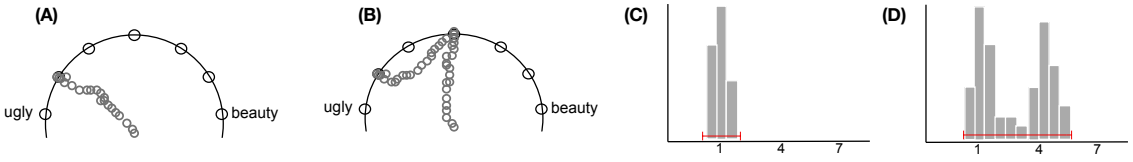



FIGURE 6.3: Mouse Tracker: empirical patterns of mouse movements (A,B) and their associated histograms (C,D). Note that the pattern (A) and (B) are different but share the same final response: in particular, (A) shows a pattern with low spatial uncertainty, by contrast (B) shows a patter with an higher level of uncertainty. The histograms (C) and (D) reflect this difference in terms of uncertainty. Interval-valued data are depicted in red color.

*Data analysis and results*. The IMedA algorithms required 455 iterations to the convergence for the $\mathcal{S}_1$ system and 671 iterations for the $\mathcal{S}_2$ system, obtaining good performances in terms of general fit ($R^2_M = 0.49$ and $R^2_Y = 0.65$). Table 6.5 shows the final estimated parameters and effects for the model. Overall, the variance explained by the pathways of the model was $\Lambda = 0.83$ whereas the single effects contributed for $38\%$, $42\%$ and $20\%$, respectively. In particulars, the centers of PQ did not show a strong impact on HQ ($\xi^c_1 = 0.10$) as well as on F ($\xi^c_2 = 0.13$). By contrast, the ranges of PQ were moderately related to HQ ($\xi^r_1 = -0.42$) and strongly to F ($\xi^c_2 = -0.84$). HQ and F positively related to B for both centers ($\gamma^c_1 = 0.48$, $\gamma^c_2 = 0.23$) and ranges ($\gamma^r_1 = 1.10$, $\gamma^r_2 = 0.49$). The direct impact of PQ on B was positive for centers ($\beta^c = 0.69$) and negative for the ranges ($\beta^r = -1.01$). As a consequence, the direct effect of PQ was mainly caused by the ranges ($\text{DE}_r = -1.01$). Interestingly, the results also showed how the indirect effects of the model were mainly generated by the ranges ($\text{IE}_r = 0.58$ and $\text{IE}_r = -0.10$). In general, the larger the perceived usability of

ESSE-3, the higher the perceived beauty of the system. Such relation become stronger with decreasing levels of uncertainty for the PQ responses. Moreover, PQ increased the perceived hedonic quality (HQ) of the system and such relation was stronger with decreasing levels of uncertainties for the PQ responses. The same pattern was also observed for PQ and F. Conversely, HQ and F positively increased the perceived beauty of ESSE-3 and, in this case, also the variability of the ranges contributed in explaining such relations. Overall, the results state that when the usability (PQ) of ESSE-3 is high, also the hedonic quality (HQ) of the system together with its beauty (B) are high. Likewise, the more the exposition to the system (F), the higher the perceived beauty (B).

| | | first mediator (HQ) | | second mediator (F) | |
|---|---|---|---|---|---|
| | | values | 95% CIs | values | 95% CIs |
| Model parameters (M) | $\mathbf{A}^c$ | 7.72 | (4.01,9.02) | -2.03 | (-0.45,3.04) |
| | $\mathbf{A}^r$ | -1.32 | (-1.59,2.46) | 1.97 | (-0.19,3.11) |
| | $\mathbf{\Pi}$ | -1.19 | (-1.88,0.05) | -1.12 | (-1.27,0.08) |
| | $\boldsymbol{\xi}^c$ | 0.10 | (0.02,0.55) | 0.13 | (0.09,0.32) |
| | $\boldsymbol{\xi}^r$ | -0.42 | (-0.77,-0.04) | -0.84 | (-1.21,-0.19) |
| Model parameters (Y) | $\alpha^c$ | 3.23 | (-0.11,4.01) | | |
| | $\alpha^l$ | 2.16 | (0.70,2.98) | | |
| | $\delta$ | 0.09 | (0.02,0.19) | | |
| | $\beta^c$ | 0.69 | (0.06,1.18) | | |
| | $\beta^r$ | -1.01 | (-1.43,-0.16) | | |
| | $\boldsymbol{\gamma}^c$ | 0.48 | (0.10,1.01) | 0.23 | (0.06,0.41) |
| | $\boldsymbol{\gamma}^r$ | 1.10 | (0.56,1.71) | 0.49 | (0.20,0.81) |
| Effects | $\mathrm{DE}_c$ | 0.69 | (0.06,1.18) | | |
| | $\mathrm{DE}_r$ | -1.01 | (-1.43,-0.16) | | |
| | $\mathrm{IE}_c$ | -0.07 | (-0.20,-0.02) | 0.06 | (-0.05,0.19) |
| | $\mathrm{IE}_r$ | 0.58 | (0.12,0.70) | -0.10 | (-0.23,-0.03) |

*Note:* $R^2_M = 0.49$, $R^2_Y = 0.65$, $\lambda_{\mathrm{DE}} = 0.38$, $\lambda^1_{\mathrm{IE}} = 0.42$, $\lambda^2_{\mathrm{IE}} = 0.20$, CIs indicate the 95% confidence intervals obtained by bias-corrected and accelerating (BCa) bootstrap with 5000 bootstrap samples.

TABLE 6.5: Second case study: Estimated model parameters and effects

## 6.8 Conclusions

In this chapter, we developed a novel and simple model (IMedA) to perform mediation analysis on interval-valued variables. As far as we know, IMedA is the first proposal that is devoted to mediation analyses of interval data. Globally, the main characteristic of this model is its use of two linear equations systems for modeling the interval pathways among the independent, mediators, and dependent observed variables. This involved the extension of the well-known Stolzenberg's decomposition to handle with interval-valued causal effects. Relatedly, a set of variance-based indices was also defined to quantify the sizes of such effects in the interval context. Finally, we used a simulation study and two real applications to highlight some characteristics of the proposed model. In particular, the simulation study revealed that the IMedA model is generally robust and accurate enough to reproduce the observed relationships among the interval variables. Moreover, our findings also showed how IMedA outperforms existing mediation approaches for single-valued variables that might be eventually used in modeling interval pathways.

### 6.8.1 Model's advantages

One nice property of the IMedA representation is that the regressions for ranges are obtained as linear combinations with the deterministic estimated centers of the intervals (*CR dependence assumption*). Consequently, in estimating the model's parameters, the IMedA algorithm works similarly to an alternating recursive two-steps procedure where the reconstruction of ranges proceeds conditionally on the reconstruction of centers. This assumption is straightforward in an interval framework, where both the centers and ranges are key components in determining the observed interval data structures. To understand this assumption, it is important to remind that our model can

be regarded as a special case of the general interval *centre-range models* introduced by Lima Neto and Carvalho (2008) where, by contrast, ranges are directly estimated from the observed data. Note that, although the latter models could be more flexible in capturing all the possible linear relationships among the observed variables, IMedA would still be able to offer a good compromise between model simplicity and model flexibility. For instance, the CR dependence assumption would possibly allow to constrain the regression parameters of ranges to be in their positive domains during the estimation procedure as well as to provide a simplified decomposition of effects on the regression equations. Moreover, the IMedA model also provides a more parsimonious representation as it requires estimating only $12k$ parameters against $16k$ of the more general *centre-range model* (with a benefit of $4k$ parameters). Finally, IMedA-ALS allows a simple generalization of the single-valued case as it subsumes the mediation model for single-valued variables as a special case. Indeed, when data are expressed in terms of degenerated intervals, the CR-parametrization always boils down to single-valued variables (with $\mathbf{x}^{\mathbf{r}} = \mathbf{y}^{\mathbf{r}} = \mathbf{0}_{n\times 1}$ and $\mathbf{M}^{\mathbf{r}} = \mathbf{0}_{n\times k}$) and the regression systems $\mathcal{S}_1$ and $\mathcal{S}_2$ simply reduce to the regressions for the ordinary multiple mediation analysis (MedA), namely $\mathbf{M} = \mathbf{1}\mathbf{A} + \mathbf{x}\xi + \mathbf{E}$ and $\mathbf{y} = \mathbf{1}\alpha + \mathbf{x}\beta + \mathbf{M}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$ (note that in this special case also the IMedA estimators as well as the effects decomposition reduce accordingly).

### 6.8.2 Model's limitations

However, as for other statistical procedures, also the proposed method can potentially suffer from some limitations. First, if the IMedA model is fitted to empirical data which are largely corrupted by noise, the corresponding estimates may violate the natural constraints of interval-valued data (namely: $\mathbf{M}^{\mathbf{r}} > \mathbf{0}_{n\times k}$ and $\mathbf{y}^{\mathbf{r}} > \mathbf{0}_{n\times 1}$) thus possibly yielding unfeasible solutions. In these situations, a constrained version of

the algorithm based on specific optimization techniques should instead be used (**lima2010constrained**). However, for the standard unconstrained algorithm, a simple way out might consist in setting to zero all the negative range estimates so that their natural constraints are numerically satisfied. Second, the *CR dependence assumption* can be a concern when more sophisticated empirical paths are considered. As a consequence, in these rare cases, the model representation can result in poor model fits in reconstructing the observed interval data structures. Third, some empirical contexts may require more complex models to better evaluate the relationships among the observed variables. For instance, we may think to moderated mediation models in which the indirect paths are partially or completely moderated by other intervening variables (e.g., age, gender, income. See: Edwards and Lambert, 2007). In a similar way, the structural assumption that mediators work in parallel could be contrived in some particular contexts where models with serial mediators would be instead preferred (Taylor, MacKinnon, and Tein, 2008). Finally, although our model considers the covariance among mediators only indirectly (i.e., without any modeling of their covariance matrix), in some circumstances researchers may prefer to adopt specific parametric covariance matrix estimators in testing mediation assumptions. In this case, because of its regression based approach, the IMedA representation appears clearly inadequate to achieve such an advantage.

### 6.8.3 Future extensions

Various possible extensions of our approach can be considered in future works. For example, moderated mediation models for interval-valued data would allow the modeling of more complex situations in which the mediated paths vary as a function of third moderator variables. Likewise, interval mediation models with serial mediators may also extend our proposal to represent situations in which researchers

evaluate three-path mediated effects. Further, also modeling correlational paths among the mediator directly, by means of parametric covariance matrix estimators, can constitute a future target. Clearly, these structural limitations could vanish in a more general interval-valued SEM approach and its development can represent a future venue of our research to be grounded on previous research in this framework (Little et al., 2007; Iacobucci, Saldanha, and Deng, 2007). Finally, several empirical situations may also require the use of mediation models handling with non-linear decomposition of effects (Hayes and Preacher, 2010). The extension of the IMedA representation to deal with non-linear pathways among variables can surely be considered an interesting future extension of the present chapter.

## Appendix A: solutions for IMedA model

$$\operatorname{vec}(\widehat{\mathbf{A}}^{r}) = (\mathbf{I}_k \otimes \mathbf{1}^T\mathbf{1})^{-1} \cdot (\mathbf{I}_k \otimes \mathbf{1})^T \operatorname{vec}(\mathbf{M^r} - \mathbf{X\Xi\Pi} - \mathbf{1A}^c\mathbf{\Pi}); \tag{A1}$$

$$\operatorname{vec}(\widehat{\mathbf{\Pi}}) = \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{\Xi}^T\mathbf{X}^T\mathbf{X\Xi}) + \\ (\mathbf{I}_k \otimes \mathbf{\Xi}^T\mathbf{X}^T\mathbf{1A}^c) + \\ (\mathbf{I}_k \otimes \mathbf{A}^{cT}\mathbf{1}^T\mathbf{1A}^c) \end{bmatrix}^{-1} \cdot \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{X\Xi})^T \operatorname{vec}(\mathbf{M^r} - \mathbf{1A^r}) + \\ (\mathbf{I}_k \otimes \mathbf{1A}^c)^T \operatorname{vec}(\mathbf{M^r} - \mathbf{1A^r}) \end{bmatrix}; \tag{A2}$$

$$\operatorname{vec}(\widehat{\mathbf{\Xi}}) = \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{X}^T\mathbf{X}) + \\ (\mathbf{\Pi} \otimes \mathbf{X}^T\mathbf{X}) \end{bmatrix}^{-1} \cdot \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{X})^T \operatorname{vec}(\mathbf{M^c} - \mathbf{1A}^c) + \\ (\mathbf{\Pi} \otimes \mathbf{X})^T \operatorname{vec}(\mathbf{M^r} - \mathbf{1A}^r - \mathbf{1A}^c\mathbf{\Pi}) \end{bmatrix}; \tag{A3}$$

$$\operatorname{vec}(\widehat{\mathbf{A}}^{c}) = \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{1}^T\mathbf{1}) + \\ (\mathbf{\Pi} \otimes \mathbf{1}^T\mathbf{1}) \end{bmatrix}^{-1} \cdot \begin{bmatrix} (\mathbf{I}_k \otimes \mathbf{1})^T \operatorname{vec}(\mathbf{M^c} - \mathbf{X\Xi}) + \\ (\mathbf{\Pi} \otimes \mathbf{1})^T \operatorname{vec}(\mathbf{M^r} - \mathbf{1A}^r - \mathbf{X\Xi\Pi}) \end{bmatrix}; \tag{A4}$$

$$\widehat{\delta} = \begin{bmatrix} \alpha^c\mathbf{1}^T\mathbf{1}\alpha^c + 2\alpha^c\mathbf{1}^T\mathbf{X}\boldsymbol{\beta} + 2\alpha^c\mathbf{1}^T\mathbf{M}\boldsymbol{\gamma} + \\ \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{1}\alpha^c + 2\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} + 2\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{M}\boldsymbol{\gamma} + \\ \boldsymbol{\gamma}^T\mathbf{M}^T\mathbf{1}\alpha^c + 2\boldsymbol{\gamma}^T\mathbf{M}^T\mathbf{X}\boldsymbol{\beta} + 2\boldsymbol{\gamma}^T\mathbf{M}^T\mathbf{M}\boldsymbol{\gamma} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \alpha^c\mathbf{1}^T(\mathbf{y}^r - \mathbf{1}\alpha^r) + \\ \boldsymbol{\beta}^T\mathbf{X}^T(\mathbf{y}^r - \mathbf{1}\alpha^r) + \\ \boldsymbol{\gamma}^T\mathbf{M}^T(\mathbf{y}^r - \mathbf{1}\alpha^r) \end{bmatrix}; \tag{A5}$$

$$\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \mathbf{X}^T\mathbf{X} + \\ \delta\mathbf{X}^T\mathbf{X}\delta \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{X}^T(\mathbf{y}^c - \mathbf{1}\alpha^c - \mathbf{M}\boldsymbol{\gamma}) + \\ \mathbf{X}^T[\,\mathbf{y}^r - \mathbf{1}\alpha^r + (-\mathbf{1}\alpha^c - \mathbf{M}\boldsymbol{\gamma})\delta\,]\delta \end{bmatrix}; \tag{A6}$$

$$\widehat{\boldsymbol{\gamma}^c} = \begin{bmatrix} \mathbf{M^c}^T\mathbf{M^c} + \\ \delta\mathbf{M^c}^T\mathbf{M^c}\delta \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{M^c}^T(\mathbf{y}^c - \mathbf{1}\alpha^c - \mathbf{X}\boldsymbol{\beta} - \mathbf{M^r}\boldsymbol{\gamma^r}) + \\ \mathbf{M^c}^T[\,\mathbf{y}^r - \mathbf{1}\alpha^r + (-\mathbf{1}\alpha^c - \mathbf{X}\boldsymbol{\beta} - \mathbf{M^r}\boldsymbol{\gamma^r})\delta\,]\delta \end{bmatrix}; \tag{A7}$$

$$\widehat{\gamma^r} = \begin{bmatrix} \mathbf{M^{r}}^T\mathbf{M^r} + \\ \delta\mathbf{M^{r}}^T\mathbf{M^r}\delta \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{M^{r}}^T(\mathbf{y}^r - \mathbf{1}\alpha^r - \mathbf{X}\boldsymbol{\beta} - \mathbf{M^c}\boldsymbol{\gamma^c}) + \\ \mathbf{M^{r}}^T[\,\mathbf{y}^r - \mathbf{1}\alpha^r + (-\mathbf{1}\alpha^r - \mathbf{X}\boldsymbol{\beta} - \mathbf{M^c}\boldsymbol{\gamma^c})\delta\,]\delta \end{bmatrix}; \tag{A8}$$

$$\widehat{\alpha}^c = \frac{1}{n(1+\delta^2)} \cdot \begin{bmatrix} \mathbf{1}^T(\mathbf{y}^c - \mathbf{X}\boldsymbol{\beta} - \mathbf{M}\boldsymbol{\gamma}) + \\ \mathbf{1}^T[\,\mathbf{y}^r - \mathbf{1}\alpha^r + (-\mathbf{X}\boldsymbol{\beta} - \mathbf{M}\boldsymbol{\gamma})\delta\,]\delta \end{bmatrix}; \tag{A9}$$

$$\widehat{\alpha}^r = \frac{1}{n} \cdot \mathbf{1}^T[\,\mathbf{y^r} - (\mathbf{1}\alpha^c - \mathbf{X}\boldsymbol{\beta} - \mathbf{M}\boldsymbol{\gamma})\delta]; \tag{A10}$$

where vec(.) is the linear operator that converts a $n \times k$ matrix into a $kn \times 1$ vector, $\otimes$ denotes the Kronecker product, $\mathbf{I}_k$ is a $k \times k$ identity matrix whereas $\mathbf{1}$ is a $n \times k$ matrix of all ones.

## Appendix B: Decomposition of effects for IMedA model

In order to derive direct and indirect effects for the IMedA model, we proceed as follows. Consider the regression system $\mathcal{S}_1$ and $\mathcal{S}_2$ shown in Eq. 6.1:

$$\mathcal{S}_1 : \begin{cases} \mathbf{M^c} = \mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi} + \mathbf{E}^c \\ \mathbf{M^r} = \mathbf{1}\mathbf{A}^r + (\mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi})\boldsymbol{\Pi} + \mathbf{E^r} \end{cases} \quad \mathcal{S}_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma^r} + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + (\mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + \mathbf{M^c}\boldsymbol{\gamma^c} + \mathbf{M^r}\boldsymbol{\gamma^r})\delta + \boldsymbol{\epsilon}^r \end{cases}$$

Firstly, substitute the equations of $\mathbf{M^c}$ and $\mathbf{M^r}$ into $\mathbf{y^c}$ and $\mathbf{y^r}$, as follows:

$$\mathcal{S'}_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + [\mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi} + \mathbf{E}^c]\boldsymbol{\gamma}^c + [\mathbf{1}\mathbf{A}^r + (\mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi})\boldsymbol{\Pi} + \mathbf{E^r}]\gamma^r + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + (\mathbf{1}\alpha^c + \mathbf{X}\boldsymbol{\beta} + [\mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi} + \mathbf{E}^c]\boldsymbol{\gamma}^c + [\mathbf{1}\mathbf{A}^r + (\mathbf{1}\mathbf{A}^c + \mathbf{X}\boldsymbol{\Xi})\boldsymbol{\Pi} + \mathbf{E^r}]\gamma^r)\delta_y^r + \boldsymbol{\epsilon}^r \end{cases} \tag{6.11}$$

Multiplying through and expanding terms, using a little algebra, we obtain the following reduced form system $\mathcal{S'}_2$:

$$\mathcal{S'}_2 : \begin{cases} \mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{1}\mathbf{A}^c\boldsymbol{\gamma}^c + \mathbf{1}\mathbf{A}^r\boldsymbol{\gamma}^r + \mathbf{1}\mathbf{A}^c\boldsymbol{\Pi}\boldsymbol{\gamma}^r + \mathbf{x^c}[\boldsymbol{\beta}^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)] \\ \quad + \mathbf{x^r}[\boldsymbol{\beta}^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)] + \mathbf{E}^c\boldsymbol{\gamma}^c + \mathbf{E}^r\boldsymbol{\gamma}^r + \boldsymbol{\epsilon}^c \\ \mathbf{y^r} = \mathbf{1}\alpha^r + \mathbf{1}\alpha^c\delta + \mathbf{1}\mathbf{A}^c\boldsymbol{\gamma}^c\delta + \mathbf{1}\mathbf{A}^r\boldsymbol{\gamma}^r\delta + \mathbf{1}\mathbf{A}^c\boldsymbol{\Pi}\boldsymbol{\gamma}^r\delta + \mathbf{x^c}[\boldsymbol{\beta}^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)]\delta \\ \quad + \mathbf{x^r}[\boldsymbol{\beta}^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)]\delta + \mathbf{E}^c\boldsymbol{\gamma}^c\delta + \mathbf{E}^r\boldsymbol{\gamma}^r\delta + \boldsymbol{\epsilon}^r \end{cases} \tag{6.12}$$

Next, taking the partial derivatives of $\mathbf{y^c}$ and $\mathbf{y^r}$ with respect to $\mathbf{x^c}$ and $\mathbf{x^r}$ we have the equations for the total effect (TE) of the model, as follows:

$$\begin{cases} \frac{\partial\,\mathbf{y^c}}{\partial\,\mathbf{x^c}} = \delta\beta^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r) & \frac{\partial\,\mathbf{y^c}}{\partial\,\mathbf{x^r}} = \delta\beta^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r) \\ \frac{\partial\,\mathbf{y^r}}{\partial\,\mathbf{x^c}} = \delta\beta^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)\delta & \frac{\partial\,\mathbf{y^r}}{\partial\,\mathbf{x^r}} = \delta\beta^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)\delta \end{cases}$$

Finally, collecting and simplifying the ensuing terms, we obtain the following equations for TE:

$$\begin{aligned}
\mathrm{TE}^{y^c} &= \beta_y^c + \beta_y^r + (\boldsymbol{\beta}_m^c \circ \boldsymbol{\gamma}_m^{c\,T})\mathbf{1}_m + (\boldsymbol{\beta}_m^c \circ \boldsymbol{\gamma}_m^{r\,T} \circ \boldsymbol{\Pi})\mathbf{1}_m + (\boldsymbol{\beta}_m^r \circ \boldsymbol{\gamma}_m^{c\,T})\mathbf{1}_m + (\boldsymbol{\beta}_m^r \circ \boldsymbol{\gamma}_m^{r\,T} \circ \boldsymbol{\Pi})\mathbf{1}_m \\
\mathrm{TE}^{y^r} &= \delta_y^r[\beta_y^c + \beta_y^r + (\boldsymbol{\beta}_m^c \circ \boldsymbol{\gamma}_m^{c\,T})\mathbf{1}_m + (\boldsymbol{\beta}_m^c \circ \boldsymbol{\gamma}_m^{r\,T} \circ \boldsymbol{\Pi})\mathbf{1}_m + (\boldsymbol{\beta}_m^r \circ \boldsymbol{\gamma}_m^{c\,T})\mathbf{1}_m + (\boldsymbol{\beta}_m^r \circ \boldsymbol{\gamma}_m^{r\,T} \circ \boldsymbol{\Pi})\mathbf{1}_m]
\end{aligned} \tag{6.13}$$

which are in the general form of $\mathrm{TE} = \mathrm{DE}_c + \mathrm{DE}_r + \mathrm{IE}_{c/c} + \mathrm{IE}_{c/r} + \mathrm{IE}_{r/c} + \mathrm{IE}_{r/r}$. Note that the equation $\mathrm{TE}^{y^r}$ for $\mathbf{y^r}$ is obtained as linear combination of $\mathrm{TE}^{y^c}$ through the parameter $\delta$.

## Appendix C: Decomposition of variance for IMedA model

Considering the reduced form system $\mathcal{S}'_2$:

$$\mathcal{S}'_2 : \begin{cases}
\mathbf{y^c} = \mathbf{1}\alpha^c + \mathbf{1}\mathbf{A}^c\boldsymbol{\gamma}^c + \mathbf{1}\mathbf{A}^r\boldsymbol{\gamma}^r + \mathbf{1}\mathbf{A}^c\boldsymbol{\Pi}\boldsymbol{\gamma}^r + \mathbf{x^c}[\boldsymbol{\beta}^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)] \\
\quad + \mathbf{x^r}[\boldsymbol{\beta}^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)] + \mathbf{E}^c\boldsymbol{\gamma}^c + \mathbf{E}^r\boldsymbol{\gamma}^r + \boldsymbol{\epsilon}^c \\
\mathbf{y^r} = \mathbf{1}\alpha^r + \mathbf{1}\alpha^c\delta + \mathbf{1}\mathbf{A}^c\boldsymbol{\gamma}^c\delta + \mathbf{1}\mathbf{A}^r\boldsymbol{\gamma}^r\delta + \mathbf{1}\mathbf{A}^c\boldsymbol{\Pi}\boldsymbol{\gamma}^r\delta + \mathbf{x^c}[\boldsymbol{\beta}^c + \boldsymbol{\xi}^c(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)]\delta \\
\quad + \mathbf{x^r}[\boldsymbol{\beta}^r + \boldsymbol{\xi}^r(\boldsymbol{\gamma}^c + \boldsymbol{\Pi}\boldsymbol{\gamma}^r)]\delta + \mathbf{E}^c\boldsymbol{\gamma}^c\delta + \mathbf{E}^r\boldsymbol{\gamma}^r\delta + \boldsymbol{\epsilon}^r
\end{cases} \tag{6.14}$$

the following identities hold:

$$\begin{aligned}
\mathrm{var}(\mathbf{y^c}) &= \mathrm{cov}(\mathbf{y^c}, \mathbf{x^c}\beta^c) + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^r}\beta^r) + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{c\,T})\mathbf{1}_k) + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{r\,T} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\
&\quad + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{c\,T})\mathbf{1}_k) + \mathrm{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{r\,T} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\
&\quad + \mathrm{cov}(\mathbf{y^c}, \mathbf{E}^c\boldsymbol{\gamma}^c) + \mathrm{cov}(\mathbf{y^c}, \mathbf{E}^r\boldsymbol{\gamma}^r) + \mathrm{cov}(\mathbf{y^c}, \boldsymbol{\epsilon}^c) \\
\mathrm{var}(\mathbf{y^r}) &= \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^c}\beta^c) + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^r}\beta^r) + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{c\,T})\mathbf{1}_k) + \\
&\quad + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{r\,T} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{c\,T})\mathbf{1}_k) + \\
&\quad + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{r\,T} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\
&\quad + \mathrm{cov}(\mathbf{y^r}, \delta\mathbf{E}^c\boldsymbol{\gamma}^c\delta) + \mathrm{cov}(\mathbf{y^r}, \mathbf{E}^r\boldsymbol{\gamma}^r\delta) + \mathrm{cov}(\mathbf{y^c}, \boldsymbol{\epsilon}^r)
\end{aligned} \tag{6.15}$$

after noticing that:

$$\begin{aligned}
&\mathrm{cov}(\mathbf{y^c}, \mathbf{1}\alpha^c + \mathbf{1}\mathbf{A}^c\boldsymbol{\gamma}^c + \mathbf{1}\mathbf{A}^r\boldsymbol{\gamma}^r + \mathbf{1}\mathbf{A}^c\boldsymbol{\Pi}\boldsymbol{\gamma}^r) = 0 \\
&\mathrm{cov}(\mathbf{y^c}, \mathbf{1}\alpha^r + \mathbf{1}\alpha^c\delta + \mathbf{1}\mathbf{A}^c\boldsymbol{\gamma}^c\delta + \mathbf{1}\mathbf{A}^r\boldsymbol{\gamma}^r\delta + \mathbf{1}\mathbf{A}^c\boldsymbol{\Pi}\boldsymbol{\gamma}^r\delta) = 0
\end{aligned}$$

where $\mathrm{cov}(.)$ and $\mathrm{var}(.)$ indicates the covariance and variance operators, $\circ$ denotes

the Hadamard product whereas $\mathbf{1}_k$ is a $k \times 1$ vector of all ones. The following properties hold:

$$\begin{bmatrix} \operatorname{cov}(\mathbf{y^c}, \mathbf{x^c}\beta^c) + \operatorname{cov}(\mathbf{y^c}, \mathbf{x^r}\beta^r) + \\ \operatorname{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \operatorname{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ \operatorname{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \operatorname{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ \operatorname{cov}(\mathbf{y^c}, \mathbf{E}^c\boldsymbol{\gamma}^c) + \operatorname{cov}(\mathbf{y^c}, \mathbf{E}^r\boldsymbol{\gamma}^r) + \operatorname{cov}(\mathbf{y^c}, \boldsymbol{\epsilon}^c) \end{bmatrix} \cdot \left[\operatorname{var}(\mathbf{y^c})\right]^{-1} = 1$$
$$\begin{bmatrix} \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^c}\beta^c) + \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^r}\beta^r) + \\ \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ \operatorname{cov}(\mathbf{y^r}, \mathbf{E}^c\boldsymbol{\gamma}^c\delta) + \operatorname{cov}(\mathbf{y^r}, \mathbf{E}^r\boldsymbol{\gamma}^r\delta) + \operatorname{cov}(\mathbf{y^r}, \boldsymbol{\epsilon}^r) \end{bmatrix} \cdot \left[\operatorname{var}(\mathbf{y^r})\right]^{-1} = 1 \tag{6.16}$$

$$\begin{bmatrix} \operatorname{cov}(\mathbf{y^c}, \mathbf{x^c}\beta^c) + \operatorname{cov}(\mathbf{y^c}, \mathbf{x^r}\beta^r) + \\ \operatorname{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \operatorname{cov}(\mathbf{y^c}, \mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ \operatorname{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \operatorname{cov}(\mathbf{y^c}, \mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ \operatorname{cov}(\mathbf{y^c}, \mathbf{E}^c\boldsymbol{\gamma}^c) + \operatorname{cov}(\mathbf{y^c}, \mathbf{E}^r\boldsymbol{\gamma}^r) \end{bmatrix} \cdot \left[\operatorname{var}(\mathbf{y^c})\right]^{-1} \approx \frac{\|\mathbf{y^{c*}} - \bar{\mathbf{y}}\mathbf{^c}\|^2}{\|\mathbf{y^c} - \bar{\mathbf{y}}\mathbf{^c}\|^2}$$
$$\begin{bmatrix} \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^c}\beta^c) + \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^r}\beta^r) + \\ \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^c}(\boldsymbol{\xi}^c \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{cT})\mathbf{1}_k) + \operatorname{cov}(\mathbf{y^r}, \delta\mathbf{x^r}(\boldsymbol{\xi}^r \circ \boldsymbol{\gamma}^{rT} \circ \boldsymbol{\Pi}^T)\mathbf{1}_k) + \\ \operatorname{cov}(\mathbf{y^r}, \mathbf{E}^c\boldsymbol{\gamma}^c\delta) + \operatorname{cov}(\mathbf{y^r}, \mathbf{E}^r\boldsymbol{\gamma}^r\delta) \end{bmatrix} \cdot \left[\operatorname{var}(\mathbf{y^r})\right]^{-1} \approx \frac{\|\mathbf{y^{r*}} - \bar{\mathbf{y}}\mathbf{^r}\|^2}{\|\mathbf{y^r} - \bar{\mathbf{y}}\mathbf{^r}\|^2} \tag{6.17}$$

where $\bar{\mathbf{y}}\mathbf{^c}$ and $\bar{\mathbf{y}}\mathbf{^r}$ are $n \times 1$ vectors containing the mean values of $\mathbf{y^c}$ and $\mathbf{y^r}$ whereas $\mathbf{y^{c*}}$ and $\mathbf{y^{r*}}$ refers to the estimated reduced equations in $\mathcal{S}'_2$ without considering the residual terms $\boldsymbol{\epsilon}^c$ and $\boldsymbol{\epsilon}^r$. Note that the terms in the right side of Eq. 6.17 denotes the variance explained by the reduced system $\mathcal{S}'_2$.

# Bibliography


Abdi, Hervé and Lynne J Williams (2010). "Principal component analysis". In: *Wiley Interdisciplinary Review: Computational Statistics* 2.4, pp. 433–459.

Aiken, Lewis R (1996). *Rating scales and checklists: Evaluating behavior, personality, and attitudes.* John Wiley & Sons.

Alwin, Duane F and Robert M Hauser (1975). "The decomposition of effects in path analysis". In: *American Sociological Review*, pp. 37–47.

Anderson, Cathy D, Jack L Warner, and Cassie C Spencer (1984). "Inflation bias in self-assessment examinations: Implications for valid employee selection." In: *Journal of Applied Psychology* 69.4, p. 574.

Anderson, NH (1981). *Foundations of information integration theory*. New York [etc.]: Academic Press.

Arndt, Stephan et al. (2000). "Charting patients' course: a comparison of statistics used to summarize patient course in longitudinal and repeated measures studies". In: *Journal of psychiatric research* 34.2, pp. 105–113.

Arnett, Jeffrey Jensen, Daniel Offer, and Mark A Fine (1997). "Reckless driving in adolescence:'State'and 'trait'factors". In: *Accident Analysis & Prevention* 29.1, pp. 57–63.

Ashforth, Blake E and Fred Mael (1989). "Social identity theory and the organization". In: *Academy of management review* 14.1, pp. 20–39.

Avanzi, Lorenzo et al. (2012). "The downside of organizational identification: Relations between identification, workaholism and well-being". In: *Work & Stress* 26.3, pp. 289–307.

Avci, E. and D. Avci (2009). "An expert system based on fuzzy entropy for automatic threshold selection in image processing". In: *Expert Systems with Applications* 36.2, pp. 3077–3085.

Azzalini, Adelchi and Adrian W Bowman (1997). "Applied smoothing techniques for data analysis". In: *Oxford Statistical Science Series, Oxford*.

Baird, John C (1984). "Information theory and information processing". In: *Information processing & management* 20.3, pp. 373–381.

Baratpour, S and A Habibi Rad (2012). "Testing goodness-of-fit for exponential distribution based on cumulative residual entropy". In: *Communications in Statistics-Theory and Methods* 41.8, pp. 1387–1396.

Barca, Laura and Giovanni Pezzulo (2012). "Unfolding visual lexical decision in time". In: *PloS one* 7.4, e35932.

Barnett, Vic and Toby Lewis (1994). *Outliers in statistical data*. Vol. 3. Wiley New York.

Baron, Reuben M and David A Kenny (1986). "The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations." In: *Journal of personality and social psychology* 51.6, p. 1173.

Benítez, Juan Manuel, Juan Carlos Martín, and Concepción Román (2007). "Using fuzzy number for measuring quality of service in the hotel industry". In: *Tourism Management* 28.2, pp. 544–555.

Berg, Robert A van den et al. (2006). "Centering, scaling, and transformations: improving the biological information content of metabolomics data". In: *BMC genomics* 7.1, p. 142.

Besserve, Michel et al. (2010). "Causal relationships between frequency bands of extracellular signals in visual cortex revealed by an information theoretic analysis". In: *Journal of computational neuroscience* 29.3, pp. 547–566.

Billard, Lynne and Edwin Diday (2002). "Symbolic regression analysis". In: *Classification, Clustering, and Data Analysis*. Springer, pp. 281–288.

— (2003). "From the statistics of data to the statistics of knowledge: symbolic data analysis". In: *Journal of the American Statistical Association* 98.462, pp. 470–487.

Bisserier, A., R. Boukezzoula, and S. Galichet (2010). "A revisited approach to linear fuzzy regression using trapezoidal fuzzy intervals". In: *Information Sciences* 180.19, pp. 3653–3673.

Biswas, Ranjit (1995). "An application of fuzzy sets in students' evaluation". In: *Fuzzy sets and systems* 74.2, pp. 187–194.

Blanco-Fernández, Angela, Ana Colubi, and Marta García-Bárzana (2013). "A set arithmetic-based linear regression model for modelling interval-valued responses through real-valued variables". In: *Information Sciences* 247, pp. 109–122.

Bliese, Paul D and Carl Andrew Castro (2000). "Role clarity, work overload and organizational support: Multilevel evidence of the importance of support". In: *Work & Stress* 14.1, pp. 65–73.

Bloch, Isabelle (1999). "On fuzzy distances and their use in image processing under imprecision". In: *Pattern Recognition* 32.11, pp. 1873–1895.

Bollen, Kenneth A and Robert Stine (1990). "Direct and indirect effects: Classical and bootstrap estimates of variability". In: *Sociological methodology* 20.1, pp. 15–140.

Bro, Rasmus and Age K Smilde (2003). "Centering and scaling in component analysis". In: *Journal of Chemometrics* 17.1, pp. 16–33.

Brown, Scott and Andrew Heathcote (2003). "Averaging learning curves across and within participants". In: *Behavior Research Methods, Instruments, & Computers* 35.1, pp. 11–21.

Bruner, Jerome S and Jerome S Bruner (2009). *Actual minds, possible worlds*. Harvard University Press.

Buckley, James J (2004). *Fuzzy statistics*. Vol. 149. Springer.

Burr, Irving W (1942). "Cumulative frequency functions". In: *The Annals of Mathematical Statistics* 13.2, pp. 215–232.

Byron, M Yu et al. (2009). "Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity". In: *Advances in neural information processing systems*, pp. 1881–1888.

Cacioppo, John T and Louis G Tassinary (1990). *Principles of psychophysiology: Physical, social, and inferential elements*. Cambridge University Press.

Calcagnì, A and L Lombardi (2014). "Dynamic Fuzzy Rating Tracker (DYFRAT): a novel methodology for modeling real-time dynamic cognitive processes in rating scales". In: *Applied soft computing* 24, pp. 948–961.

Calcagnì, Antonio, Luigi Lombardi, and Eduardo Pascali (2013). "Non-convex fuzzy data and fuzzy statistics: a first descriptive approach to data analysis". In: *Soft Computing*, pp. 1–14.

Cazes, P et al. (1997). "Entension de l'analyse en composantes principales à des données de type intervalle". In: *Revue de Statistique appliquée* 45.3, pp. 5–24.

Celmiņš, A. (1987). "Least squares model fitting to fuzzy vector data". In: *Fuzzy sets and systems* 22.3, pp. 245–269.

Chan, LK, HP Kao, and ML Wu (1999). "Rating the importance of customer needs in quality function deployment by fuzzy and entropy methods". In: *International Journal of Production Research* 37.11, pp. 2499–2518.

Chang, Yu-Hern and Chung-Hsing Yeh (2002). "A survey analysis of service quality for domestic airlines". In: *European Journal of Operational Research* 139.1, pp. 166–177.

Chen, Mon Chu, John R Anderson, and Myeong Ho Sohn (2001). "What can a mouse cursor tell us more?: correlation of eye/mouse movements on web browsing". In: *CHI'01 extended abstracts on Human factors in computing systems*. ACM, pp. 281–282.

Chen, Vivien YC et al. (2011). "Fuzzy MCDM approach for selecting the best environment-watershed plan". In: *Applied Soft Computing* 11.1, pp. 265–275.

Cheng, Heng-Da and Jim-Rong Chen (1997). "Automatically determine the membership function based on the maximum entropy principle". In: *Information Sciences* 96.3, pp. 163–182.

Chung, Hsing-Tai and Daniel G Schwartz (1995). "A resolution-based system for symbolic approximate reasoning". In: *International Journal of Approximate Reasoning* 13.3, pp. 201–246.

Ciavolino, E. and J.J. Dahlgaard (2009). "Simultaneous Equation Model based on the generalized maximum entropy for studying the effect of management factors on enterprise performance". In: *Journal of applied statistics* 36.7, pp. 801–815.

Ciavolino, Enrico, Sergio Salvatore, and Antonio Calcagnì (2013). "A fuzzy set theory based computational model to represent the quality of inter-rater agreement". In: *Quality & Quantity*, pp. 1–16.

CINECA and Kion (1999). *ESSE3: Student management system*. URL: `http://www.kion.it/en/solutions/esse3-student-management-system`.

Claessens, Brigitte JC et al. (2004). "Planning behavior and perceived control of time at work". In: *Journal of Organizational Behavior* 25.8, pp. 937–950.

Colubi, Ana et al. (2001). "On the formalization of fuzzy random variables". In: *Information Sciences* 133.1, pp. 3–6.

Coppersmith, Don and Ravi Kumar (2004). "An improved data stream algorithm for frequency moments". In: *Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, pp. 151–156.

Coppi, Renato, Maria A Gil, and Henk AL Kiers (2006). "The fuzzy approach to statistical analysis". In: *Computational statistics & data analysis* 51.1, pp. 1–14.

Coppi, Renato, Paolo Giordani, and Pierpaolo D'Urso (2006). "Component models for fuzzy data". In: *Psychometrika* 71.4, pp. 733–761.

Coppi, Renato et al. (2006). "Least squares estimation of a linear regression model with LR fuzzy response". In: *Computational Statistics & Data Analysis* 51.1, pp. 267–286. ISSN: 01679473.

Corning, Peter A and Stephen Jay Kline (1998). "Thermodynamics, information and life revisited, Part I:'To be or entropy'". In: *Systems Research and Behavioral Science* 15.4, pp. 273–295.

Costa, Antonio José Leal et al. (2004). "Perception of disability in a public health perspective: a model based on fuzzy logic". In: *International journal of medical informatics* 73.7, pp. 647–656.

Costas, Concepción San Luis, Pedro Prieto Maranon, and Juan A Hernandez Cabrera (1994). "Application of diffuse measurement to the evaluation of psychological structures". In: *Quality & quantity* 28.3, pp. 305–313.

Dai, Jianhua, Wentao Wang, and Ju-Sheng Mi (2013). "Uncertainty measurement for interval-valued information systems". In: *Information Sciences* 251, pp. 63–78.

Dale, Rick, Nicholas C Hindy, and Michael J Spivey (2006). "Feature-semantic gradients in lexical categorization revealed by graded manual responses". In: *Proceedings of the 28th Annual Meeting of the Cognitive Science Society: Lawrence Erlbaum*.

Dale, Rick, Caitlin Kehoe, and Michael J Spivey (2007). "Graded motor responses in the time course of categorizing atypical exemplars". In: *Memory & Cognition* 35.1, pp. 15–28.

De Luca, Aldo and Settimo Termini (1972). "A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory". In: *Information and control* 20.4, pp. 301–312.

Deery, Hamish A (2000). "Hazard and risk perception among young novice drivers". In: *Journal of Safety Research* 30.4, pp. 225–236.

DeJoy, David M (1992). "An examination of gender differences in traffic accident risk perception". In: *Accident Analysis & Prevention* 24.3, pp. 237–246.

Delorme, Arnaud and Scott Makeig (2004). "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis". In: *Journal of neuroscience methods* 134.1, pp. 9–21.

Di Crescenzo, Antonio and Maria Longobardi (2009). "On cumulative entropies". In: *Journal of Statistical Planning and Inference* 139.12, pp. 4072–4087.

Diamond, P. (1988). "Fuzzy least squares". In: *Information Sciences* 46.3, pp. 141–157.

Diday, Edwin, Monique Noirhomme-Fraiture, et al. (2008). *Symbolic data analysis and the SODAS software*. Wiley Online Library.

Domino, George Ed and Marla L Domino (2006). *Psychological testing: An introduction*. Cambridge University Press.

Douzal-Chouakria, Ahlame, Lynne Billard, and Edwin Diday (2011). "Principal component analysis for interval-valued observations". In: *Statistical Analysis and Data Mining* 4.2, pp. 229–246.

Dubois, D. and H.M. Prade (2000). *Fundamentals of fuzzy sets*. Vol. 7. Springer.

Dubois, Didier et al. (1988). *Possibility theory: an approach to computerized processing of uncertainty*. Vol. 2. Plenum press New York.

D'Urso, Pierpaolo (2003). "Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data". In: *Computational Statistics & Data Analysis* 42.1, pp. 47–72.

D'Urso, Pierpaolo and Tommaso Gastaldi (2000). "A least-squares approach to fuzzy linear regression analysis". In: *Computational Statistics & Data Analysis* 34.4, pp. 427–440.

Eddy, J Mark and Patricia Chamberlain (2000). "Family managment and deviant peer association as mediators of the impact of treatment condition on youth antisocial behavior." In: *Journal of Consulting and Clinical Psychology* 68.5, p. 857.

Edwards, Jeffrey R and Lisa Schurer Lambert (2007). "Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis." In: *Psychological methods* 12.1, p. 1.

Elliott, Digby, Werner F Helsen, and Romeo Chua (2001). "A century later: Woodworth's (1899) two-component model of goal-directed aiming." In: *Psychological bulletin* 127.3, p. 342.

Endres, Dominik et al. (2010). "Modelling spike trains and extracting response latency with Bayesian binning". In: *Journal of Physiology-Paris* 104.3, pp. 128–136.

Engelbrecht, Sascha E (2001). "Minimum principles in motor control". In: *Journal of Mathematical Psychology* 45.3, pp. 497–542.

Eriksson, Lennart (2006). *Multi-and megavariate data analysis*. MKS Umetrics AB.

Everitt, Brian and Andrew Pickles (2004). *Statistical Aspects of the Design and Analysis of Clinical Trials*. Imperial College Press.

Everitt, BS (1995). "The analysis of repeated measures: a practical review with examples". In: *The Statistician*, pp. 113–135.

Facchinetti, G. and N. Pacchiarotti (2006). "Evaluations of fuzzy quantities". In: *Fuzzy Sets and Systems* 157.7, pp. 892–903.

Fairchild, Amanda J et al. (2009). "R2 effect-size measures for mediation analysis". In: *Behavior research methods* 41.2, pp. 486–498.

Faulkenberry, Thomas J (2014). "Hand movements reflect competitive processing in numerical cognition." In: *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* 68.3, p. 147.

Fekedulegn, Desta B et al. (2007). "Area under the curve and other summary indicators of repeated waking cortisol measurements". In: *Psychosomatic medicine* 69.7, pp. 651–659.

Ferson, Scott et al. (2004). "Summary from the epistemic uncertainty workshop: consensus amid diversity". In: *Reliability Engineering & System Safety* 85.1, pp. 355–369.

Fields, Gary S (2003). "Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the United States". In: *Research in labor economics* 22, pp. 1–38.

Finn, Peter and Barry WE Bragg (1986). "Perception of the risk of an accident by young and older drivers". In: *Accident Analysis & Prevention* 18.4, pp. 289–298.

Fischer, Martin H and Matthias Hartmann (2014). "Pushing forward in embodied cognition: may we mouse the mathematical mind?" In: *Frontiers in psychology* 5.

Fishburn, Peter C (1973). "Interval representations for interval orders and semiorders". In: *Journal of Mathematical Psychology* 10.1, pp. 91–105.

Fisher, Cynthia D and March L To (2012). "Using experience sampling methodology in organizational behavior". In: *Journal of Organizational Behavior* 33.7, pp. 865–877.

Fisher, Nicholas I (1995). *Statistical analysis of circular data*. Cambridge University Press.

Flash, Tamar and Ealan Henis (1991). "Arm trajectory modifications during reaching towards visual targets". In: *Journal of cognitive Neuroscience* 3.3, pp. 220–230.

Franco-Watkins, Ana M and Joseph G Johnson (2011). "Decision moving window: Using interactive eye tracking to examine decision processes". In: *Behavior research methods* 43.3, pp. 853–863.

Frank, Stefan L (2013). "Uncertainty reduction as a measure of cognitive load in sentence comprehension". In: *Topics in cognitive science* 5.3, pp. 475–494.

Freeman, Jonathan B and Nalini Ambady (2010). "MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method". In: *Behavior Research Methods* 42.1, pp. 226–241.

Freeman, Jonathan B, Rick Dale, and Thomas A Farmer (2011). "Hand in motion reveals mind in motion". In: *Frontiers in Psychology* 2.

Freeman, Jonathan B et al. (2008). "Will a category cue attract you? Motor output reveals dynamic competition across person construal." In: *Journal of Experimental Psychology: General* 137.4, p. 673.

Friedman, Jason, Scott Brown, and Matthew Finkbeiner (2013). "Linking cognitive and reaching trajectories via intermittent movement control". In: *Journal of Mathematical Psychology* 57.3, pp. 140–151.

Frison, Lars and Stuart J Pocock (1992). "Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design". In: *Statistics in medicine* 11.13, pp. 1685–1704.

Furnham, Adrian (1986). "Response bias, social desirability and dissimulation". In: *Personality and individual differences* 7.3, pp. 385–400.

Furnham, Adrian and Monika Henderson (1982). "The good, the bad and the mad: Response bias in self-report measures". In: *Personality and Individual Differences* 3.3, pp. 311–320.

Gallop, Robert et al. (2009). "Mediation analysis with principal stratification". In: *Statistics in medicine* 28.7, p. 1108.

Garibaldi, J.M. and R.I. John (2003). "Choosing membership functions of linguistic terms". In: *Fuzzy Systems, 2003. FUZZ'03. The 12th IEEE International Conference on*. Vol. 1. IEEE, pp. 578–583.

Garibaldi, J.M. et al. (2004). "A case study to illustrate the use of non-convex membership functions for linguistic terms". In: *Fuzzy Systems, 2004. Proceedings. 2004 IEEE International Conference on*. Vol. 3. IEEE, pp. 1403–1408.

Gil, María Ángeles and Gil González-Rodríguez (2012). "Fuzzy vs. Likert Scale in Statistics". In: *Combining Experimentation and Theory*. Springer, pp. 407–420.

Gil, María Ángeles, Miguel López-Díaz, and Dan A Ralescu (2006). "Overview on the development of fuzzy random variables". In: *Fuzzy sets and systems* 157.19, pp. 2546–2557.

Gill, Philip E, Walter Murray, and Margaret H Wright (1981). *Practical optimization*. Academic-Press.

Giordani, Paolo (2010). "Three-way analysis of imprecise data". In: *Journal of Multivariate Analysis* 101.3, pp. 568–582.

Giordani, Paolo and Henk AL Kiers (2004a). "Principal component analysis of symmetric fuzzy data". In: *Computational Statistics & Data Analysis* 45.3, pp. 519–548.

— (2004b). "Three-way component analysis of interval-valued data". In: *Journal of chemometrics* 18.5, pp. 253–264.

— (2007). "Principal component analysis with boundary constraints". In: *Journal of Chemometrics* 21.12, pp. 547–556.

Glaholt, Mackenzie G and Eyal M Reingold (2011). "Eye movement monitoring as a process tracing methodology in decision making research." In: *Journal of Neuroscience, Psychology, and Economics* 4.2, p. 125.

Glendon, A Ian et al. (1996). "Age and gender differences in perceived accident likelihood and driver competences". In: *Risk Analysis* 16.6, pp. 755–762.

Göb, Rainer, Christopher McCollin, and Maria Fernanda Ramalhoto (2007). "Ordinal methodology in the analysis of Likert scales". In: *Quality & Quantity* 41.5, pp. 601–626.

Golan, A. and G. Judge (1996). *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. Wiley.

Golfried, Marvin R and Ronald N Kent (1972). "Traditional versus behavioral personality assessment: A comparison of methodological and theoretical assumptions." In: *Psychological Bulletin* 77.6, p. 409.

Greene, Joshua and Jonathan Haidt (2002). "How (and where) does moral judgment work?" In: *Trends in cognitive sciences* 6.12, pp. 517–523.

Grubbs, Frank E (1969). "Procedures for detecting outlying observations in samples". In: *Technometrics* 11.1, pp. 1–21.

Guo, Wensheng (2004). "Functional data analysis in longitudinal settings using smoothing splines". In: *Statistical methods in medical research* 13.1, pp. 49–62.

Haidt, Jonathan (2001). "The emotional dog and its rational tail: a social intuitionist approach to moral judgment." In: *Psychological review* 108.4, p. 814.

Halff, Henry M, Andrew Ortony, and Richard C Anderson (1976). "A context-sensitive representation of word meanings". In: *Memory & Cognition* 4.4, pp. 378–383.

Hall, Harold V, Jane S Thompson, and Joseph G Poirier (2007). "Detecting deception in neuropsychological cases: Toward an applied model". In: *Forensic Examiner* 16.3, p. 7.

Hall Jr, Marshall (1948). "Distinct representatives of subsets". In: *Bulletin of the American Mathematical Society* 54.10, pp. 922–926.

Hamborg, Kai-Christoph, Julia Hülsmann, and Kai Kaspar (2014). "The interplay between usability and aesthetics: more evidence for the "what is usable is beautiful" notion". In: *Advances in Human-Computer Interaction* 2014, p. 15.

Hanss, M. (2005). *Applied fuzzy arithmetic*. Springer.

Hartmann, Donald P, Billy A Barrios, and David D Wood (2004). "Principles of behavioral observation". In: *Comprehensive handbook of psychological assessment* 3, pp. 108–127.

Hassenzahl, Marc (2004). "The interplay of beauty, goodness, and usability in interactive products". In: *Human-Computer Interaction* 19.4, pp. 319–349.

— (2008). "Aesthetics in interactive products correlates and consequences of beauty". In: *Product Experience*. Elsevier, pp. 287–302.

Hassenzahl, Marc, Michael Burmester, and Franz Koller (2003). "AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität". In: *Mensch & Computer 2003*. Springer, pp. 187–196.

Hassenzahl, Marc and Andrew Monk (2010). "The inference of perceived usability from beauty". In: *Human–Computer Interaction* 25.3, pp. 235–260.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2001). *The elements of statistical learning theory*. Springer.

Hayes, Andrew F and Kristopher J Preacher (2010). "Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear". In: *Multivariate Behavioral Research* 45.4, pp. 627–660.

Hehman, Eric, Ryan M Stolier, and Jonathan B Freeman (2014). "Advanced mouse-tracking analytic techniques for enhancing psychological science". In: *Group Processes & Intergroup Relations*, p. 1368430214538325.

Hesketh, Beryl et al. (1988). "Practical applications and psychometric evaluation of a computerised fuzzy graphic rating scale". In: *Advances in Psychology* 56, pp. 425–454.

Hesketh, Beryl et al. (1989). "Fuzzy logic: Toward measuring Gottfredson's concept of occupational social space." In: *Journal of Counseling Psychology* 36.1, pp. 103–09.

Hesketh, Tim and Beryl Hesketh (1994). "Computerized fuzzy ratings: The concept of a fuzzy class". In: *Behavior Research Methods, Instruments, & Computers* 26.3, pp. 272–277.

Hesketh, Tim, Robert Pryor, and Beryl Hesketh (1988). "An application of a computerized fuzzy graphic rating scale to the psychological measurement of individual differences". In: *International journal of man-machine studies* 29.1, pp. 21–35.

Hirsh, Jacob B, Raymond A Mar, and Jordan B Peterson (2012). "Psychological entropy: a framework for understanding uncertainty-related anxiety." In: *Psychological review* 119.2, p. 304.

Hoffmann, Errol R (2013). "Which version/variation of Fitts' law? A critique of information-theory models". In: *Journal of motor behavior* 45.3, pp. 205–215.

Howard, George S and Patrick R Dailey (1979). "Response-shift bias: A source of contamination of self-report measures." In: *Journal of Applied Psychology* 64.2, p. 144.

Huynh, Van-Nam, Tu Bao Ho, and Yoshiteru Nakamori (2002). "A parametric representation of linguistic hedges in Zadeh's fuzzy logic". In: *International Journal of Approximate Reasoning* 30.3, pp. 203–223.

Hwang, F et al. (2005). "A submovement analysis of cursor trajectories". In: *Behaviour & Information Technology* 24.3, pp. 205–217.

Iacobucci, Dawn, Neela Saldanha, and Xiaoyan Deng (2007). "A meditation on mediation: Evidence that structural equations models perform better than regressions". In: *Journal of Consumer Psychology* 17.2, pp. 139–153.

Imai, Kosuke, Luke Keele, and Dustin Tingley (2010). *A general approach to causal mediation analysis.*

Imai, Kosuke, Luke Keele, and Teppei Yamamoto (2010). "Identification, inference and sensitivity analysis for causal mediation effects". In: *Statistical Science*, pp. 51–71.

Imai, Kosuke and David A Van Dyk (2004). "Causal inference with general treatment regimes". In: *Journal of the American Statistical Association* 99.467.

Irpino, Antonio (2006). "Spaghetti PCA analysis: An extension of principal components analysis to time dependent interval data". In: *Pattern recognition letters* 27.5, pp. 504–513.

Irpino, Antonio, Carlo Lauro, and Rosanna Verde (2003). "Visualizing symbolic data by closed shapes". In: *Between Data Science and Applied Data Analysis*. Springer, pp. 244–251.

Jahanshahloo, Gholam Reza, F Hosseinzadeh Lotfi, and Mohammad Izadikhah (2006). "Extension of the TOPSIS method for decision-making problems with fuzzy data". In: *Applied Mathematics and Computation* 181.2, pp. 1544–1551.

Jansen, Anthony R, Alan F Blackwell, and Kim Marriott (2003). "A tool for tracking visual attention: The Restricted Focus Viewer". In: *Behavior research methods, instruments, & computers* 35.1, pp. 57–69.

Johnson, Addie et al. (2012). "Action as a Window to Perception: Measuring Attention with Mouse Movements". In: *Applied Cognitive Psychology* 26.5, pp. 802–809.

Jonah, Brian A (1986). "Accident risk and risk-taking behaviour among young drivers". In: *Accident Analysis & Prevention* 18.4, pp. 255–271.

Judd, Charles M and David A Kenny (1981). "Process analysis estimating mediation in treatment evaluations". In: *Evaluation review* 5.5, pp. 602–619.

Kacprzyk, J. and M. Fedrizzi (1992). *Fuzzy regression analysis*. Vol. 1. Physica-Verlag.

Kaiser, Henry F (1958). "The varimax criterion for analytic rotation in factor analysis". In: *Psychometrika* 23.3, pp. 187–200.

Kao, Chiang and S-Tai Liu (2000). "Data envelopment analysis with missing data: an application to university libraries in Taiwan". In: *Journal of the Operational Research Society*, pp. 897–905.

Kaufmann, Arnold and DL Swanson (1975). *Introduction to the theory of fuzzy subsets*. Vol. 1. Academic Press New York.

Kiers, Henk AL (1997). "Techniques for rotating two or more loading matrices to optimal agreement and simple structure: A comparison and some technical details". In: *Psychometrika* 62.4, pp. 545–568.

— (2002). "Setting up alternating least squares and iterative majorization algorithms for solving various matrix optimization problems". In: *Computational statistics & data analysis* 41.1, pp. 157–170.

— (2004). "Bootstrap confidence intervals for three-way methods". In: *Journal of Chemometrics* 18.1, pp. 22–36.

Kiers, Henk AL and Jos MF ten Berge (1989). "Alternating least squares algorithms for simultaneous components analysis with equal component weight matrices in two or more populations". In: *Psychometrika* 54.3, pp. 467–473.

Kirkup, Les and Robert B Frenkel (2006). *An Introduction to Uncertainty in Measurement: Using the GUM (Guide to the Expression of Uncertainty in Measurement)*. Cambridge University Press.

Koop, Gregory J (2013). "An assessment of the temporal dynamics of moral decisions". In: *Judgment and Decision Making* 8.5, pp. 527–539.

Kristal, Alan R et al. (2000). "Mediating factors in dietary change: understanding the impact of a worksite nutrition intervention". In: *Health Education & Behavior* 27.1, pp. 112–125.

Kumaresan, Ramdas, Donald W Tufts, and Louis L Scharf (1984). "A Prony method for noisy data: Choosing the signal components and selecting the order in exponential signal models". In: *Proceedings of the IEEE* 72.2, pp. 230–233.

Lai, Shih-Chiung, Gottfried Mayer-Kress, and Karl M Newell (2006). "Information entropy and the variability of space-time movement error". In: *Journal of motor behavior* 38.6, pp. 451–466.

Lai, Shih-Chiung et al. (2005). "Information entropy analysis of discrete aiming movements". In: *Acta psychologica* 119.3, pp. 283–304.

Lalla, Michelle, Gisella Facchinetti, and Giovanni Mastroleo (2005). "Ordinal scales and fuzzy set systems to measure agreement: an application to the evaluation of teaching activity". In: *Quality and Quantity* 38.5, pp. 577–601.

Laming, Donald (2010). "Statistical information and uncertainty: A critique of applications in experimental psychology". In: *Entropy* 12.4, pp. 720–771.

Lauro, Carlo N and Francesco Palumbo (2000). "Principal component analysis of interval data: a symbolic data analysis approach". In: *Computational Statistics* 15.1, pp. 73–87.

Le-Rademacher, Jennifer and Lynne Billard (2012). "Symbolic Covariance Principal Component Analysis and Visualization for Interval-Valued Data". In: *Journal of Computational and Graphical Statistics* 21.2, pp. 413–432.

Lee, S., S. Kim, and N.Y. Jang (2008). "Design of Fuzzy Entropy for Non Convex Membership Function". In: *Advanced Intelligent Computing Theories and Applications. With Aspects of Contemporary Intelligent Computing Techniques*, pp. 55–60.

Li, Linyi and Deren Li (2008). "Fuzzy entropy image segmentation based on particle swarm optimization". In: *Progress in Natural Science* 18.9, pp. 1167–1171.

Li, Qing (2012). "A novel Likert scale based on fuzzy sets theory". In: *Expert Systems with Applications*.

Lima Neto, Eufrásio and Francisco De Carvalho (2010). "Constrained linear regression models for symbolic interval-valued variables". In: *Computational Statistics & Data Analysis* 54.2, pp. 333–347. ISSN: 01679473.

Lima Neto, Eufrásio de A and Francisco de AT de Carvalho (2008). "Centre and Range method for fitting a linear regression model to symbolic interval data". In: *Computational Statistics & Data Analysis* 52.3, pp. 1500–1515.

Link, Stephen, James T Townsend, and F Gregory Ashby (1985). *The Stochastic Modeling of Elementary Psychological Processes*.

Little, Roderick JA (1993). "Statistical analysis of masked data". In: *Journal of Official Statistics - Stockholm* 9, pp. 407–407.

Little, Todd D et al. (2007). "Structural equation modeling of mediation and moderation with contextual factors". In: *Modeling contextual effects in longitudinal studies*. Lawrence Erlbaum Associates, pp. 207–230.

Liu, G (2002). "On velocity estimation using position measurements". In: *American Control Conference, 2002. Proceedings of the 2002*. Vol. 2. IEEE, pp. 1115–1120.

Luce, R Duncan (1956). "Semiorders and a theory of utility discrimination". In: *Econometrica, Journal of the Econometric Society*, pp. 178–191.

— (1996). "The ongoing dialog between empirical science and measurement theory". In: *journal of mathematical psychology* 40.1, pp. 78–98.

— (2003). "Whatever happened to information theory in psychology?" In: *Review of general psychology* 7.2, p. 183.

Lucero, Jorge C and Laura L Koenig (2000). "Time normalization of voice signals using functional data analysis". In: *The Journal of the Acoustical Society of America* 108.4, pp. 1408–1420.

MacKinnon, David (2008). *Introduction to statistical mediation analysis*. Routledge.

MacKinnon, David P and Amanda J Fairchild (2009). "Current directions in mediation analysis". In: *Current Directions in Psychological Science* 18.1, pp. 16–20.

MacKinnon, David P et al. (2002). "A comparison of methods to test mediation and other intervening variable effects." In: *Psychological methods* 7.1, p. 83.

Magnuson, James S (2005). "Moving hand reveals dynamics of thought". In: *Proceedings of the National Academy of Sciences of the United States of America* 102.29, pp. 9995–9996.

Matthews, JN et al. (1990). "Analysis of serial measurements in medical research." In: *BMJ: British Medical Journal* 300.6719, p. 230.

McGuire, Jonathan et al. (2009). "A reanalysis of the personal/impersonal distinction in moral psychology research". In: *Journal of Experimental Social Psychology* 45.3, pp. 577–580.

Medasani, Swarup, Jaeseok Kim, and Raghu Krishnapuram (1998). "An overview of membership function generation techniques for pattern recognition". In: *International Journal of approximate reasoning* 19.3, pp. 391–417.

Meredith, William and Roger E Millsap (1985). "On component analyses". In: *Psychometrika* 50.4, pp. 495–507.

Meyer, David E et al. (1988). "Optimality in human motor performance: ideal control of rapid aimed movements." In: *Psychological review* 95.3, p. 340.

Michell, Joel (1999). *Measurement in psychology: A critical history of a methodological concept*. Vol. 53. Cambridge University Press.

Miller, Delbert C and Neil J Salkind (2002). *Handbook of research design and social measurement*. Sage.

Millsap, Roger E and William Meredith (1988). "Component analysis in cross-sectional and longitudinal data". In: *Psychometrika* 53.1, pp. 123–134.

Mood, AM and FA Graybill (1974). *Introduction to the theory of statistics*.

Moon, Chiung, Joosung Lee, and Siyeong Lim (2010). "A performance appraisal and promotion ranking system based on fuzzy logic: An implementation case in military organizations". In: *Applied Soft Computing* 10.2, pp. 512–519.

Moore, Ramon E (1966). *Interval analysis*. Vol. 4. Prentice-Hall Englewood Cliffs.

Morein-Zamir, Sharon et al. (2006). "Measuring online volitional response control with a continuous tracking task". In: *Behavior research methods* 38.4, pp. 638–647.

Morett, Laura M and Brian Macwhinney (2013). "Syntactic transfer in English-speaking Spanish learners". In: *Bilingualism: Language and Cognition* 16.01, pp. 132–151.

Mueller, Florian and Andrea Lockerd (2001). "Cheese: tracking mouse movement activity on websites, a tool for user modeling". In: *CHI'01 extended abstracts on Human factors in computing systems*. ACM, pp. 279–280.

Navarro, Jorge, Yolanda del Aguila, and Majid Asadi (2010). "Some new results on the cumulative residual entropy". In: *Journal of Statistical Planning and Inference* 140.1, pp. 310–322.

Nguyen, Hung T and Berlin Wu (2006). *Fundamentals of statistics with fuzzy data*. Springer.

Nichols, Shaun and Ron Mallon (2006). "Moral dilemmas and moral rules". In: *Cognition* 100.3, pp. 530–542.

Nieradka, Grzegorz and Bohdan Butkiewicz (2007). "A method for automatic membership function estimation based on fuzzy measures". In: *Foundations of Fuzzy Logic and Soft Computing*. Springer, pp. 451–460.

Nkurunziza, Sévérien and S Ejaz Ahmed (2011). "Estimation strategies for the regression coefficient parameter matrix in multivariate multiple regression". In: *Statistica Neerlandica* 65.4, pp. 387–406.

Nunnally, Jum (1978). *Psychometric methods*.

OECD (2011). *Employment rate*. URL: `http://www.oecd.org/statistics/`.

— (2013). *Unemployment rates*. URL: `http://www.oecd.org/statistics/`.

Ogden, Jane (2007). *Essential readings in health psychology*. McGraw-Hill Education (UK).

Omar, Rumana Z et al. (1999). "Analysing repeated measurements data: a practical comparison of methods". In: *Statistics in medicine* 18.13, pp. 1587–1603.

O'Reilly, Christian and Réjean Plamondon (2011). "Can computer mice be used as low-cost devices for the acquisition of planar human movement velocity signals?" In: *Behavior Research Methods* 43.1, pp. 229–238.

Palumbo, Francesco and Carlo N Lauro (2003). "A PCA for interval-valued data based on midpoints and radii". In: *New developments in Psychometrics*. Springer, pp. 641–648.

Park, Sangun, Murali Rao, and Dong Wan Shin (2012). "On cumulative residual Kullback–Leibler information". In: *Statistics & Probability Letters* 82.11, pp. 2025–2032.

Pashler, Hal and John Wixted (2002). *Stevens' Handbook of Experimental Psychology, Methodology in Experimental Psychology*. Vol. 4. Wiley.

Pearl, Judea et al. (2009). "Causal inference in statistics: An overview". In: *Statistics Surveys* 3, pp. 96–146.

Pettit, Frances Annie (2002). "A comparison of World-Wide Web and paper-and-pencil personality questionnaires". In: *Behavior Research Methods, Instruments, & Computers* 34.1, pp. 50–54.

Pew, RW and DA Rosenbaum (1988). "Human movement control: Computation, representation, and implementation". In: *Stevens' handbook of experimental psychology* 2, pp. 473–509.

Plamondon, Réjean (1995). "A kinematic theory of rapid human movements". In: *Biological cybernetics* 72.4, pp. 295–307.

Plamondon, Réjean and Adel M Alimi (1997). "Speed/accuracy trade-offs in target-directed movements". In: *Behavioral and Brain Sciences* 20.02, pp. 279–303.

Poli, Riccardo, James Kennedy, and Tim Blackwell (2007). "Particle swarm optimization". In: *Swarm intelligence* 1.1, pp. 33–57.

Prado Martín, Fermín Moscoso del (2011). "Macroscopic thermodynamics of reaction times". In: *Journal of Mathematical Psychology* 55.4, pp. 302–319.

Preacher, Kristopher J and Andrew F Hayes (2008). "Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models". In: *Behavior research methods* 40.3, pp. 879–891.

Preacher, Kristopher J and Ken Kelley (2011). "Effect size measures for mediation models: quantitative strategies for communicating indirect effects." In: *Psychological methods* 16.2, p. 93.

Preacher, Kristopher J, Derek D Rucker, and Andrew F Hayes (2007). "Addressing moderated mediation hypotheses: Theory, methods, and prescriptions". In: *Multivariate behavioral research* 42.1, pp. 185–227.

Rai, Tage S and Keith J Holyoak (2010). "Moral principles or consumer preferences? Alternative framings of the trolley problem". In: *Cognitive Science* 34.2, pp. 311–321.

Ramsay, James O (2006). *Functional data analysis*. Wiley Online Library.

Ramsay, James O and Bernard W Silverman (2002). *Applied functional data analysis: methods and case studies*. Vol. 77. Springer New York.

Rao, Murali et al. (2004). "Cumulative residual entropy: a new measure of information". In: *Information Theory, IEEE Transactions on* 50.6, pp. 1220–1228.

Ratcliff, Roger (1998). "The role of mathematical psychology in experimental psychology". In: *Australian journal of psychology* 50.3, pp. 129–130.

Reuter, U. (2008). "Application of non-convex fuzzy variables to fuzzy structural analysis". In: *Soft Methods for Handling Variability and Imprecision*, pp. 369–375.

Richman, Joshua S and J Randall Moorman (2000). "Physiological time-series analysis using approximate entropy and sample entropy". In: *American Journal of Physiology-Heart and Circulatory Physiology* 278.6, H2039–H2049.

Rio, Alejandro Quintela-del and Graciela Estevez-Perez (2012). "Nonparametric kernel distribution function estimation with kerdiest: an R package for bandwidth choice and applications". In: *Journal of Statistical Software* 50.8, pp. 1–21.

Rosenbaum, David A (2005). "The Cinderella of psychology: the neglect of motor control in the science of mental life and behavior." In: *American Psychologist* 60.4, p. 308.

Ross, Timothy J (2009). *Fuzzy logic with engineering applications*. Wiley.

Rosseel, Yves (2012). "Lavaan: An R Package for Structural Equation Modeling". In: *Journal of Statistical Software* 48.2, pp. 1–36. URL: `http://www.jstatsoft.org/v48/i02/`.

Rosso, Osvaldo A (2007). "Entropy changes in brain function". In: *International Journal of Psychophysiology* 64.1, pp. 75–80.

Rowe, William D (1994). "Understanding uncertainty". In: *Risk analysis* 14.5, pp. 743–750.

Roychowdhury, Shounak and Witold Pedrycz (2001). "A survey of defuzzification strategies". In: *International Journal of intelligent systems* 16.6, pp. 679–695.

Sáa, Sara de la Rosa de et al. (2013). "Fuzzy Rating vs. Fuzzy Conversion Scales: An Empirical Comparison through the MSE". In: *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*. Springer, pp. 135–143.

Saal, Frank E., Ronald G. Downey, and Mary Anne Lahey (1980). "Rating the ratings: assessing the psychometrics quality of rating data". In: *Psychological Bulletin* 88.2, pp. 413–428.

Salicone, Simona (2007). *Measurement uncertainty: An approach via the mathematical theory of evidence*. Springer.

Sawyer, John E (1992). "Goal and process clarity: Specification of multiple constructs of role ambiguity and a structural equation model of their antecedents and consequences." In: *Journal of Applied Psychology* 77.2, p. 130.

Seibold, David R and Robert D McPhee (1979). "Commonality analysis: A method for decomposing explained variance in multiple regression analyses". In: *Human Communication Research* 5.4, pp. 355–365.

Selen, Luc PJ, Michael N Shadlen, and Daniel M Wolpert (2012). "Deliberation in the motor system: reflex gains track evolving evidence leading to a decision". In: *The Journal of Neuroscience* 32.7, pp. 2276–2286.

Senn, Stephen, Lynda Stevens, and Nish Chaturvedi (2000). "Repeated measures in clinical trials: simple strategies for analysis using summary measures". In: *Statistics in medicine* 19.6, pp. 861–877.

Seow, Steven C (2005). "Information theoretic models of HCI: a comparison of the Hick-Hyman law and Fitts' law". In: *Human-Computer Interaction* 20.3, pp. 315–352.

Shadmehr, Reza (2005). *The computational neurobiology of reaching and pointing: a foundation for motor learning*. MIT press.

Shannon, Claude (1956). "The bandwagon". In: *IRE Transactions on Information Theory* 1.2, p. 3.

Sijtsma, Klaas (2012). "Psychological measurement between physics and statistics". In: *Theory & Psychology*, pp. 786–809.

Smith, Eliot R and Jamie DeCoster (2000). "Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems". In: *Personality and social psychology review* 4.2, pp. 108–131.

Smits-Engelsman, BCM, GP Van Galen, and Jaak Duysens (2002). "The breakdown of Fitts' law in rapid, reciprocal aiming movements". In: *Experimental Brain Research* 145.2, pp. 222–230.

Sobel, Michael E (1982). "Asymptotic confidence intervals for indirect effects in structural equation models". In: *Sociological methodology* 13.1982, pp. 290–312.

— (1990). "Effect analysis and causation in linear structural equation models". In: *Psychometrika* 55.3, pp. 495–515.

Song, Joo-Hyun and Ken Nakayama (2009). "Hidden cognitive states revealed in choice reaching tasks". In: *Trends in cognitive sciences* 13.8, pp. 360–366.

Spivey, Michael J and Rick Dale (2006). "Continuous dynamics in real-time cognition". In: *Current Directions in Psychological Science* 15.5, pp. 207–211.

Stephen, Damian G, James A Dixon, and Robert W Isenhower (2009). "Dynamics of representational change: entropy, action, and cognition." In: *Journal of Experimental Psychology: Human Perception and Performance* 35.6, p. 1811.

Stöber, Joachim and Jutta Joormann (2001). "A short form of the Worry Domains Questionnaire: Construction and factorial validation". In: *Personality and individual differences* 31.4, pp. 591–598.

Stolzenberg, Ross M (1980). "The measurement and decomposition of causal effects in nonlinear and nonadditive models". In: *Sociological methodology*, pp. 459–488.

Taheri, S Mahmoud (2003). "Trends in fuzzy statistics". In: *Austrian journal of statistics* 32.3, pp. 239–257.

Tang, Rong and Hans-Georg Müller (2008). "Pairwise curve synchronization for functional data". In: *Biometrika*.

Taris, Toon W, Annet H de Lange, and Michiel AJ Kompier (2010). "Research methods in occupational health psychology". In: *Leka, S.; Houdmont, J.(ed.), Occupational Health Psychology*, pp. 269–297.

Taubman-Ben-Ari, Orit, Mario Mikulincer, and Amit Iram (2004). "A multi-factorial framework for understanding reckless driving—appraisal indicators and perceived environmental determinants". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 7.6, pp. 333–349.

Taylor, Aaron B, David P MacKinnon, and Jenn-Yun Tein (2008). "Tests of the three-path mediated effect". In: *Organizational Research Methods* 11.2, pp. 241–269.

Taylor, Shelley E (1991). "Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis." In: *Psychological bulletin* 110.1, p. 67.

Thompson, Shelby G et al. (2007). "Kinematic analysis of multiple constraints on a pointing task". In: *Human movement science* 26.1, pp. 11–26.

Timmerman, Marieke E and Henk AL Kiers (2002). "Three-way component analysis with smoothness constraints". In: *Computational statistics & data analysis* 40.3, pp. 447–470.

Toderi, Stefano et al. (2013). "Psychometric properties of the UK and Italian versions of the HSE stress indicator tool". In: *European Journal of Psychological Assessment* 29.1, pp. 72–79.

Tourangeau, Roger and Kenneth A Rasinski (1988). "Cognitive processes underlying context effects in attitude measurement." In: *Psychological bulletin* 103.3, p. 299.

Tractinsky, Noam, Adi Shoval Katz, and Dror Ikar (2000). "What is beautiful is usable". In: *Interacting with computers* 13.2, pp. 127–145.

Trefethen, Lloyd N and David Bau (1997). *Numerical linear algebra*. 50. Siam.

Trevino, Linda Klebe (1986). "Ethical decision making in organizations: A person-situation interactionist model". In: *Academy of management Review* 11.3, pp. 601–617.

Vanderweele, Tyler and Stijn Vansteelandt (2009). "Conceptual issues concerning mediation, interventions and composition". In: *Statistics and its Interface* 2, pp. 457–468.

Verkuilen, J. and M. Smithson (2006). *Fuzzy set theory: Applications in the social sciences*. Vol. 147. Sage Publications, Incorporated.

Viertl, Reinhard (1996). "Real data and their mathematical description for parameter estimation". In: *Mathematical Modelling of Systems* 2.4, pp. 262–298.

— (2006). "Univariate statistical analysis with fuzzy data". In: *Computational Statistics & Data Analysis* 51.1, pp. 133–147.

— (2011). *Statistical methods for fuzzy data*. Wiley.

Voss, Andreas and Jochen Voss (2008). "A fast numerical algorithm for the estimation of diffusion model parameters". In: *Journal of Mathematical Psychology* 52.1, pp. 1–9.

Vrij, Aldert and Bruno Verschuere (2013). *Lie detection in a forensic context*. Oxford bibliographies in psychology. New York, NY: Oxford University Press.

Walczyk, Jeffrey J et al. (2005). "LYING PERSON-TO-PERSON ABOUT LIFE EVENTS: A COGNITIVE FRAMEWORK FOR LIE DETECTION". In: *Personnel Psychology* 58.1, pp. 141–170.

Walker, Neff, David E Meyer, and John B Smelcer (1993). "Spatial and temporal characteristics of rapid cursor-positioning movements with electromechanical mice in human-computer interaction". In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 35.3, pp. 431–458.

Wang, Fei and Baba C Vemuri (2007). "Non-rigid multi-modal image registration using cross-cumulative residual entropy". In: *International journal of computer vision* 74.2, pp. 201–215.

Wang, Fei et al. (2003). "Cumulative residual entropy, a new measure of information & its application to image alignment". In: *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, pp. 548–553.

Wel, RP van der et al. (2009). "Trajectories emerging from discrete versus continuous processing models in phonological competitor tasks: a commentary on Spivey, Grosjean, and Knoblich (2005)." In: *Journal of experimental psychology. Human perception and performance* 35.2, p. 588.

White, Corey N et al. (2010). "Using diffusion models to understand clinical disorders". In: *Journal of Mathematical Psychology* 54.1, pp. 39–52.

Windschitl, Paul D and Gary L Wells (1996). "Measuring psychological uncertainty: Verbal versus numeric methods". In: *Journal of Experimental Psychology Applied* 2, pp. 343–364.

Wu, Wei et al. (2006). "Bayesian population decoding of motor cortical activity using a Kalman filter". In: *Neural computation* 18.1, pp. 80–118.

Yang, Miin-Shen and Cheng-Hsiu Ko (1996). "On a class of fuzzy c-numbers clustering procedures for fuzzy data". In: *Fuzzy sets and systems* 84.1, pp. 49–60.

Yuan, Ke-Hai, Ying Cheng, and Scott Maxwell (2013). "Moderation Analysis Using a Two-Level Regression Model". In: *Psychometrika*, pp. 1–32.

Zhang, Zhiyong and Lijuan Wang (2013). "Methods for mediation analysis with missing data". In: *Psychometrika* 78.1, pp. 154–184.

Ziegler, Matthias, Carolyn MacCann, and Richard Roberts (2011). *New perspectives on faking in personality assessment*. Oxford University Press, USA.

Zimmermann, Hans Jürgen (2001). *Fuzzy set theory-and its applications*. Springer.