Università degli Studi di Padova
Facoltà di Psicologia

Corso di Laurea Specialistica in Psicologia Clinica

A Bayesian Approach to Cognitive algebra. Methodology for Model Selection in Cognitive Psychology.

Relatore
Ch.mo Prof. Giulio Vidotto

Laureando
Marco Vicentini

Matricola
529637-SP1C

Anno Accademico 2005–2006

Sommario

La teoria dell'Integrazione dell'Informazione di N. H. Anderson presenta differenti modelli cognitivi di integrazione delle informazioni, siano esse percettive, psicofisiche, come pure attributive o di giudizio.
L'algebra cognitiva rappresenta lo strumento teorico capace di esprimere i differenti processi cognitivi di integrazione attraverso una rappresentazione matematica. L'espressione algebrica di tale processi permette di verificare la bontà dei differenti modelli. Anderson individua tra i principali processi integrativi quelli additivi, moltiplicativi e di media ponderata, esprimendoli con adeguate formulazioni algebriche (cap. 1).
Particolare attenzione viene dedicata a quest'ultimo processo cognitivo, algebricamente non lineare, che esprime i processi cognitivi di integrazione tramite la coppia di parametri Importanza $\times$ Valore (cap. 2).
Nel testo vengono indicati differenti approcci metodologici che permettono di verificare la capacità dei modelli di spiegare validamente i dati. Inoltre viene introdotto il principio metodologico della semplicità, espresso tramite un approccio bayesiano, al fine di formulare ed implementare un algoritmo capace di selezionare il modello ottimale tra i diversi modelli concorrenti. (cap. 3).
I dati provenienti da due differenti esperimenti, uno di psicofisica, inerente la fisica ingenua, l'altro legato ai giudizi di fiducia, vengono utilizzati per

esemplificare la metodologia precedentemente indicata, utilizzando la funzione R-AVERAGE appositamente scritta in R (cap. 4). Infine vengono presentati i risultati di un preliminare confronto tra i differenti algoritmi di stima dei pesi e dei valori per il modello "averaging".

Introduction

The integration of the elements of a complex source is a general question which crosses the domains of science. Actually, the general framework of the Information Integration Theory (IIT) is developed only in the single domains of science.

In neuroscience, Tononi (2004) proposes the IIT of consciousness, which is defined as the capacity of a system to integrate information. This theory claims that the informational relationships among the elements of a complex determine the quality of consciousness. These relationships are specified by the values of effective information among them. This theory accounts for several neurobiological observations concerning consciousness.
In computational and information sciences, IIT provides some uniform query interfaces to heterogeneous information sources; it is based on the algebraic theory of incomplete information (Arens, Knoblock, & Shen, 1996; Ullman, 2000; Grahne & Kiricenko, 2004). This theory postulates a global schema which provides a unifying data model for all the information sources.
In psychology, it is difficult to find a general integration theory, due to the micro-theories which characterise the contemporary psychology (Noble & Shanteau, 1999).

Anderson (1981, 1982) lays out his theory of information integration in cognitive psychology. His approach covers a large variety of psychological fields, such as psychophysics, memory, cognitive development, social development, and language processing. The keys to IIT can be found in the functional perspective, cognitive algebra, and functional measurement theory.

Functional perspective is based on the purposefulness of thought and action, which are conceptualised in terms of their functions in a goal-directed behaviour and can be captured by a value. The measurement of the value is necessary to determine the goal objective.
Anderson develops the cognitive algebra which connects the internal, subjective variables to the overt stimuli and behaviours. He suggests that, whether the internal variables are integrated by some algebraic rules, the pattern of responses can be used to diagnose the form of those rules. In fact, analysing the data graphs, there are distinct patterns which imply one of three general algebraic rules. A pattern of parallelism implies the use of addition rules, a linear fan pattern implies the use of multiplicative rules, and a crossover pattern implies the use of an averaging rule.
Chapter 1 and 2 provide an introduction to the just mentioned basic rationale of IIT and to its derived scaling methodology, that is the functional measurement.

Two general problems in cognitive algebra are dealt with in chapter 3. One is a problem of model diagnosis, that it to distinguish among the integration rules (Singh & Bhargava, 1986). The adding, averaging and multiplying models can similarly account for the response variability, but not with the

same efficiency. A first approach to the model selection is qualitative, observing the factorial graph. More evidences can be provided by suitable factorial designs and by formal statistical analyses. The analysis of variance may suggest the integration rule. But only a theoretical framework for the model selection, as the Bayesian approach, can supply some operative criteria to solve the model uncertainty. The Bayesian methodological approach provides the capability to select the optimal model.
A second problem is specifically concerned with the numerically estimation and comparison of the parameters for the averaging model, which is inherently non-linear (Zalinski, 1987). Different procedures are presented and a new one, the R-AVERAGE function, is implemented. This algorithm allows to compare the progressive and to validate the assumptions which characterise the averaging model.

In chapter 4 two different areas of study are chosen to examine and to evaluate solutions to the problems outlined above. These experiments aim at showing the functional measurement methodology applied to assess the physical knowledge and personal judgements. The suitable methodology allows for the model selection of designs with two and more factors.
Lastly, we carry out a comparison between two different algorithms which allow to estimate the weight and value parameters for the averaging model.

Key-words Information Integration Theory, Cognitive Algebra, Bayesian Model Selection, R project.

Acknowledgements

Many people helped me during the years I spent in Padua.

No line of this thesis could have been written without my parents, Pietro and Giuliana, and without my cherished Paola, who constantly taught me what is really important.

Many people at the Department of General Psychology also helped and encouraged me in different ways: Alessia, Luca, Gianmarco, and all the tyros I worked with in the Methodology Laboratory. In particular, my colleague and friend Andrea.

Finally, I would like to thank prof. Giulio Vidotto for his encouragement and guidance.

Contents

Chapter 1

Information integration theory

The theory of information integration, originally proposed by Anderson (1962a, 1974a, 1974b, 1974c), aims to develop a unified, general theory of everyday cognition. It deals with two issues: multiple determination and personal value. Within the cognitive psychology, integration theory answers to these two requests in the form of algebraic integration schemas. These schemas provide the capability to measure any personal value.

1.1 Unified theory

The theory of information integration represents a unified, general theory. In the last forty years, its generality appeared in different psychological areas, covering psychophysics, functional memory, language processing, cognitive development, judgement decision, moral judgement and social cognition. Its unity appeared in the applicability of the same concepts and methods in all these domains (Anderson, 1981, 1982).
Two axioms underlie the theory: purposiveness and integration.

- Purposiveness may be considered an axiom of psychology, since thought and action are basically goal oriented. Psychology requires a functional perspective which conceptualises any biological and social goals in terms of purposiveness. This last one requires a one-dimensional representation of thought and action, manifested in the approach-and-avoidance values. Turning value into a scientific working key, implies a theory of measurement for psychological values. Such a functional measurement theory was developed with the cognitive algebra.

- Integration explains that perception, thought and action depend on multiple determinants. They operate always, in any biological or social interaction. Each of them is assessed by a value, positive or negative, with reference to the goal. These multiple values are then integrated to obtain an overall net value, which governs goal-directed action.

Information integration theory has been developed to offer a solution to the two axioms, measuring the values of separate stimuli on true psychological scales and finding the law which governs the integration of these separate values (Anderson, 1996, 2001b, 2004).

Integration theory is based on four interlocking concepts: *stimulus integration*, *stimulus valuation*, *cognitive algebra*, *functional measurement*.

Valuation The chain of processing which transforms the physical stimulus into its psychological counterpart is represented by the valuation operation. Valuation refers to the processes which extract the information from observable physical stimuli, which can be potentially controlled in experimental studies. The task instructions set some dimensions of judgement. That is, each stimulus is assessed by some values. This value may be an immediate sensory effect, as for example a sound, or a semantic inference, as a word. These scale values are not enough. A concept of weight is also necessary

for many integration tasks. The weight represents the relative salience or importance of each stimulus in the whole response. So, a representation of the stimulus requires two parameters, that is weight and value.
Integration theory is primarily concerned with stimuli at the psychological level, because they are the immediate causes of thought and behaviour. In general, the weight and value representation depends very sensitively on the prevailing dimension of judgement and also on the momentary motivational state of the organism. The concept of valuation takes account of the fundamental importance of representing individual differences within the theory.

Integration Integration theory is concerned with the study of stimulus integration, studying how they are combined, and analysing the effective stimuli. Virtually, every thought and behaviour is multiply caused, it is the result of numerous co-acting factors, and the joint action of multiple stimuli. Single causes are seldom sufficient to understand or predict. In everyday life, the multiple causation is the rule.
Multiple causation may be examined from two related points of view: synthesis and analysis. Synthesis studies the response to a complex stimulus field, perceptual as well as social. It corresponds to the integration function that represents how the effective stimuli combine to produce the response. Analysis is inverse to synthesis, and tries to dissect a given response into its causal components.
When several factors are involved, each of them pushing in its own directions, their combined effect is not generally predictable without the aid of quantitative analysis, generally in terms of psychological values of the individual. Without such quantitative capability, many basic problems of multiple causation can hardly be touched.

Functional measurement contributes to the analysis since it dissects the observed response into its functional components. The efficacy of this approach is connected to the fact that stimulus integration often obeys algebraic models. These ones are sufficiently common to indicate the existence of a general cognitive algebra of multiple causation.

Cognitive algebra Writers as far back as Aristotle, in his Nicomachean Ethics, conjectured that human judgement obeys algebraic rules in various situations. But these ones remained conjectures; they could not be tested without psychological measurement.

The measurement problem was solved with the methodology of functional measurement. The essential idea of functional measurement is to establish the algebraic rule as a simultaneous solution for all the unobservable factors. The algebraic integration rules provide metric variables and structures for the measurement of those variables (Anderson, 1962a; Anderson & Zalinski, 1990).

Functional measurement diagram Figure 1.1 shows how valuation, integration and cognitive algebra are interlocked in a joint solution. Physical

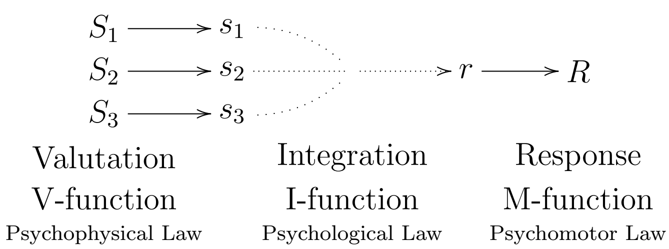


Figure 1.1: Functional measurement diagram. Chain of three linked functions go from observable stimulus field to observable response. The valuation function maps physical stimuli into subjective counterparts. The integration function maps the subjective stimulus field into an implicit response. The response function maps the implicit response into an observable response.

stimuli, S, have an impact on the organism and are processed by the valuation function $\mathbb{V}$ into their psychological values, s. These psychological stimuli are combined by the integration function, $\mathbb{I}$, into an implicit response, r. This one is externalised by the response function, M, to become the observable response R. The path from the observable stimulus, S, to the observable response, R, is represented by three linked functions. These are:

$$
\begin{aligned}
&\text{Valuation function:} && \mathbb{V}(S) = s; \\
&\text{Integration function:} && \mathbb{I}(s) = r; \\
&\text{Response function:} && \mathbb{M}(r) = R.
\end{aligned}
$$

The observable stimuli and response are denoted by the uppercase letters, S and R, whereas the lowercase ones, s and r, are used to indicate their unobservable, subjective counterparts. As a solution to the problem of measuring the psychological values of the stimuli, the functional measurement proposes to measure the psychological value of the response and to determine the psychological law or integration function, $\mathbb{I}$ (Anderson, 1990a).

Functional measurement Functional measurement provides a unification of ideas and methods which constitutes a general theory of psychological measurement. The fundamental element is the integration function, $\mathbb{I}$. Its mathematical form carries implicit scales of stimulus and response variables. This functional form provides the structural frame of the scale and its validational base. The term "functional measurement" derives from this fundamental property of the integration function. That is, the stimulus values are those that are functional in the thoughts and behaviours under study.
Implicit in the notion of cognitive algebra is a numeric representation of the stimuli. Say that two stimuli are averaged or multiplied seems to presuppose numerical values. Accordingly, the study of any algebraic rule is integrally bound up with the measurement of psychological values.

The guiding principle of functional measurement is that measurement scales derive from the substantive theory. "The logic of the present scaling technique consists in using the postulated behaviour laws to induce a scaling on the dependent variable" (Anderson, 1962b, p. 410). In terms of the diagram of Figure 1.1, behaviour law corresponds to the psychological law or integration function, $\mathbb{I}$. Dependent variables refer to the overt response, R, which is transformed into a linear scale; that is, a linear function of the underlying response, r.

A key problem is the development of procedures that could ensure a valid linear response scale. A general theory of measurement must be able to work with monotone (ordinal) response scales. Observed response measures, in general, will not be linear, but they will often be monotone functions of the underlying response variable. Since the observed response is a monotone scale, some monotone transformation will make it a linear scale. If the integration function is valid, then the desired transformation can be computed because it is the one that makes the data fit the function form.

Since the integration function depends on two or more variables, it allows the determination of the monotone transformation and still leaves degrees of freedom to test whether the transformed data fit the function. If the postulated integration function is not valid, then the data will not in general pass this test. This test of goodness of fit is essential; it provides the validational criterion of the integration function and of the derived scales. Only if algebraic models of stimulus integration are empirically valid this approach can have a meaningful value.

Many technical problems arise in implementing monotone analysis. These problems can be greatly simplified if the observed response scale can be linearized by experimental procedures rather than by statistical computation. That is why numerical response scales have been emphasised in the experimental work.

1.2 Judgement decision in multi-attribute evaluation

The trade-off is a characteristic of choice and judgement decisions, both in everyday life and in the major social and economic decisions. Political compromise itself is a form of trade-off.
It is possible to deal with the complexity of the choice among several alternatives with a simple solution: by representing each alternative with its values for each of the several attributes, or dimensions, such as price or quality, and by weighting each value with the importance of the corresponding attribute. The complexity of the valuation is reduced by dealing with one attribute at a time. The complexity of the integration of these values is reduced by applying a mechanical formula, called the weighted sum of values. Choosing the best alternative becomes merely choosing the highest weighted sum. This seemingly effective technique is called multi-attribute evaluation (Edwards & Newman, 1982). Multiattribute evaluation is an optimal rule, providing a powerful tool for decision analysis.

Beyond its simplicity, multi-attribute analysis presents a fundamental difficulty concerning the measurement. The application of the multi-attribute formula depends critically on the measurement of weights and values. Unless these are valid measures, the choice prescribed by the formula may be far from optimal. In fact, the multiattribute formula requires strong measurement assumptions, which can be easily violated. Values and weights must be on linear scales, and the weights have a known zero. Even more stringent is the assumption of the common unit for these scales, which is essential to add up the attribute values. There may be violations of these measurement assumptions, so that the multiattribute formula can erroneously assign the highest value to a less preferred alternative (Oral & Kettani, 1989; Pöyhönen & Hämäläinen, 2001).

Most of the applications involve subjective values, which lie outside any normative multiattribute framework. Disordinality may result from measurement biases. Multiattribute analysis aims to put the best alternative in first place. Unless the measured values and weights are veridical, the multiattribute formula may be incorrect, putting a less desirable alternative in first place.
The resolution of the measurement issue requires the development of a self-estimation methodology, in which judges estimate directly the weights and values of several separate attributes. Therefore, in order to avoid measurement biases, it is required a theory of psychological measurement (Anderson, 1996, cap. 13).
The approach of integration information theory approach seeks to determine the cognitive integration operators. These operators provide the base and frame for the measurement of values and weights. The measurement is basically cognitive, even in multiattribute analysis; values are, in fact, typically personal and subjective.

A primary problem in the measurement of values is to obtain a linear scale in which the observed values are a linear function of the underlying preferences. Another problem is that weights are confounded with the unit of the value scale, and so are not generally identifiable. The multiattribute formula has a further requirement: different attributes must be measured in a common currency. The common unit requirement stems from the reliance on separate measures of weight and value, essential in judgement decision or multiattribute analysis. The functional measurement methodology seems to provide linear equal-interval scales that are needed for multiattribute analysis (Anderson & Zalinski, 1990).

The rating method can provide linear measures of values, where the observable response, R, is a linear function of the unobservable, r. The rating method has provided an efficient solution to the problem of fundamental measurement of subjective, psychological variables. Only few simple precautions, mainly end-anchors and preliminary practice (Anderson, 1982, sect. 6.2) are needed to avoid certain biases, setting up a stable frame of reference, where ratings are as a true linear scale.

The self-estimation methodology can be put on a solid foundation with functional measurement. The functional measures constitute a validity criterion for the self-estimated measures. With a validity criterion, current methods of self-estimation can be improved or discarded. If it is used an appropriate factorial design, it would be obtained the overall integrated judgement of each combination of the stimulus attributes, and the self-estimates of the weight-value parameters of each separate attribute.
Functional measurement can provide separate measure of individual attributes as well as valid scales of weights and values of the stimulus variables. This provides a validational criterion for improving the method of self-estimation,

especially in the fields in which no simple algebraic rule applies.

1.3 Integration models

Many works showed that psychophysics, judgement-decision theory, learning-motivation, social-cognition and developmental psychology often follow simple algebraic models (Anderson & Cuneo, 1978a; Anderson, 1989). In many different areas of psychology, the human organism frequently appears to average, sum or multiply the stimulus information in order to arrive at a response. There is a psychophysical law which connects psychological sensation to the physical stimulus, providing a practicable method for simultaneous measurement of the subjective probability and utility. Various studies in psycholinguists, person perception and decision theory also found use for algebraic models. These algebraic rules are generically defined cognitive algebra (Anderson & Cuneo, 1978a).
This theme is illustrated by the *parallelism theorem*, according to which an addition rule will produce in the factorial plot a pattern of parallelism, and by the *linear fan theorem*, which says that a multiplication rule will produce a linear fan pattern (Anderson, 1981).

1.3.1 Parallelism theorem

Addition can be conceptualised as a stepwise movement along a response continuum. At each successive step, the last response is adjusted moving sideways an amount equal to the value of the present stimulus, positive or negative. This integration process requires a minimal cognitive capacity. The addition rule needs that the value of any stimulus is independent from the amount of the prior information (Anderson, 1996, pp. 65–66).

The essential idea of the parallelism analysis is simple. It is possible to test the hypothesis that two or more stimulus variables add together to yield the observed response. If the hypothesis is true, manipulating the stimulus variables in a factorial design, the factorial plot of the response data will exhibit a pattern of parallelism, sign of an adding-type operation.

Given two physical stimuli denoted by S_{Ai} and S_{Bj}, it is possible to indicate with s_{Ai} and s_{Bj} the subjective values of the physical stimuli. The experimental conditions are pairs of physical stimuli, (S_{Ai}, S_{Bj}). In the adding model, the subject's implicit response is assumed to be a sum of the subjective values of the given stimuli. The implicit value of the overt response is denoted by r_{ij}. So, the adding model may be written as

$$r_{ij} = s_{Ai} + s_{Bj} \tag{1.1}$$

and the observable response, R_{ij} is on a linear equal-interval scale, so that

$$R_{ij} = c_0 + c_1 r_{ij} \tag{1.2}$$

where c_0 and c_1 are zero and unit constants. In terms of the functional measurement diagram of figure 1.1, the essential assumption is that integration function, $\mathbb{I}$, is additive. Thus, the model assumes additivity at the subjective level.

If the adding model of equation 1.1 is true, and if the observable response is a linear equal-interval scale, then the factorial data plot will form a set of parallel curves with no interaction. Moreover, the row means of the factorial design will be estimate as the subjective values of the row stimuli on validated equal-interval scales and the same for the column means (Anderson, 1981, p. 15). Algebraically, the entries in every row have a constant difference in every column. Geometrically, means from every row of data will plot as parallel curves, as shown in figure 1.2.

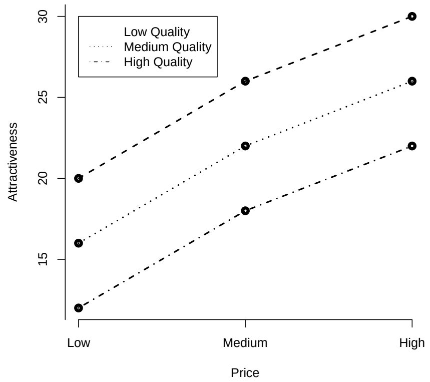


Figure 1.2: Parallelism theorem.
Hypothetical data illustrate parallelism analysis. The three curves represent attractiveness as a function of price for different levels of quality. The parallelism in the graph shows an addition schema.

The parallelism theorem provides a remarkably simple and precise way to test the model. If either assumption is incorrect, then the parallelism will not in general be obtained. There is, of course, a logical possibility that non linearity in the response scale balances non additivity in the integration rule to yield net parallelism. However, the observed parallelism supports the adding model (equation 1.1), the linearity of the response scale (equation 1.2), providing linear scales of the stimulus variables.
The analysis of variance provides exact tests of goodness of fit. If all variables are integrated by adding-type operations, then interactions are zero in principle and are expected to be non significant in practice.

Within functional measurement theory, both measurement problems, of response and of stimulus, are treated as organic components of the substantive rule of stimulus integration. Similar analyses are supported by algebraic rules. Mathematically, the parallelism theorem is elementary. The most difficult part is to establish it empirically. The fact is that deviations from parallelism are not infrequent; most of these deviations derives from the ubiquitous averaging process, which yields the non-parallelism with differential weighting. Many works showed the operation of a general cognitive algebra of judgement-decision.

1.3.2 Linear fan theorem

Multiplication rules seem natural in many areas of psychology: the motivation seems to act as an energiser of ability in the determination of the performance; the expectancy of success appears to act as a proportionality coefficient on the value of the goal; language quantifiers, such as "very", seem to operate as multipliers. Multiplication can be performed as a fractional process. In order to judge expected value of a single probabilistic outcome, the outcome is located on the response continuum according to its full value. The probability fractionate this location.

Interesting complications appear in "as-if" multiplication, when linear fan patterns can appear without any kind of multiplication.

A multiplication formula implicitly suggests that its terms correspond to cognitive entities. The analysis of cognitive units requires the study of integration processes. The study of cognitive algebra, accordingly, requires methods for testing and analysing multiplication rules (Anderson, 1996, pp. 66–67).

A multiplying rule of stimulus integration can be diagnosed by a linear fan pattern. Suppose that a multiplying model holds, so that

$$r_{ij} = s_{Ai} \times s_{Bj} \tag{1.3}$$

and the observed response, R_{ij} is on a linear equal-interval scale, so that

$$R_{ij} = c_0 + c_1 r_{ij} \tag{1.4}$$

Then the appropriate factorial graph will form a fan of straight lines, as shown in figure 1.3. The factorial graph requires that the column stimuli be spaced on the horizontal axis according to their subjective values. Thus, the linear fan pattern will only be obtained if the factorial graph is constructed appropriately. Moreover the row means of the factorial design will be estimate the subjective values of the row stimuli on linear scales, and similarly for the column means.

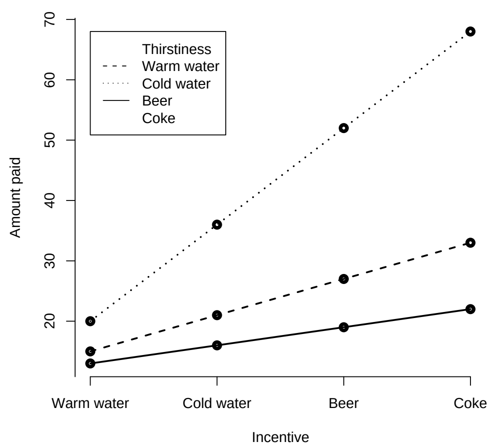


Figure 1.3: Linear fan theorem.
Hypothetical data illustrate multiplying analysis. The three curves represent amount paid for each drink under different levels of motivation and incentive. The linear fan in the graph shows multiplying schema.

By the same logic as the parallelism theorem, an observed linear fan support both assumptions of the theorem. This supports the multiplying model (equation 1.3), the linearity of the response measure (equation 1.4), providing linear scales of the subjective stimulus variables.

The graphical test will need a supplemental statistical test, as suggest also by Masin (2004). The linear regression analysis can provide this capability. The regular Row × Column interaction term is split into two components: the linear × linear and the residual. The linear × linear component represents the linear fan pattern; the residual represents a deviation from the linear fan. A complete test requires a significant linear × linear component and a non significant residual. Details are given in Anderson (2001b, pp. 259–279, 485–505).

1.3.3 Paradoxical non-additivity

A positive experience may actually decrease the net affective state; in many personal and social schema, adding positive information can have negative effects, as shown recently by Girard, Mullet, and Callahan (2002) and by Falconi and Mullet (2003). This paradoxical finding casts doubt on any kind of adding or multiplying rule. It seems to raise doubt about any simple linear rule of integration.
Thought and action were found to obey an averaging rule in many tasks in which the addition rule and the multiplying rule have failed. The same positive informer could have incremental or decremental effects, depending on what the other informer was integrated with.

According to Anderson (1990a), in virtually every domain of psychology there is a cross-over curve that rules out the adding or the multiplying model. In the adding model, adding an item of positive value should increase the overall judgement regardless of the original information. That is, $R_2 = w_1 s_1 + w_2 s_2$, is always greater than $R_1 = w_1 s_1$ if $w_2 s_2$ is positive (where R_n is the overall judgement based on n pieces of information, s_i, $i = 1$ to n, and the slope of R_2 as a function of s_1 is always w_1, the same as the slope of R_1). Hence, $R_2(s_1)$ is always parallel to $R_1(s_1)$ for any given value of s_2.
When parallelism is obtained, the interpretation is reasonably straightforward. When parallelism is not obtained, the interpretation is difficult. The deviation from parallelism could have been produced by non-linear biases in the response, by non-linearity in the integration operation, or by violations

of the independence assumption of no stimulus interaction.
"No general rule can be given for the interpretation of deviations from parallelism" (Anderson, 1981, p. 21). Empirical contributes and integration experiments from psychophysic, psycholinguist, developmental psychology, comparison processes, and social schema report non-parallel observed pattern. The presence of the interaction, as shown in figure 1.4, rules out the strict adding model.

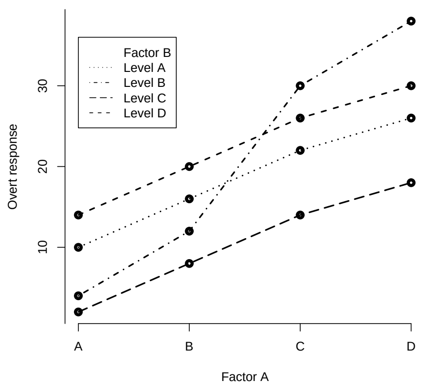


Figure 1.4: Crossover interaction.
Typical pattern of crossover interaction reported in studies supporting the averaging model.

In 1965 Anderson started off with a more general model than equation 1.2, predicting the parallelism as well as the crossover interaction; this could be written as

$$R_{ij} = c_0 + w_{Ai}s_{Ai} + w_{Bj}s_{Bj} \tag{1.5}$$

where w_{Ai} and w_{Bj} are the weights or importance parameters. This model is capable to predict that $R_2(s_1)$ and $R_1(s_1)$ may not be always parallel. As shown in chapter 2, the crossover interaction requires the consideration of an essentially new rule of integration, given by the averaging model.

1.3.4 Generalised algebraic models

The adding and multiplying models are easily generalised in order to allow more stimulus variables and also to let both adding and multiplying operations within one model. These models may be studied experimentally by treating each stimulus variable as a design factor.
In a multi-factor adding model the parallelism theorem, denoted by equation 1.2, may be applied directly to each and every pair of factors. Each two-way factorial graph should be a set of parallel curves. Treating each serial position as a design factor, the model analysis can dissect the response into its separate serial components.
In a similar way, linear fan analysis, expressed by equation 1.4, can be generalised to handle more factors. Each and every pair of factors in a multi-factor multiplying model should exhibit the linear fan pattern. Three factor multiplying models arise occasionally, but no four factor multiplying model is known. There is some evidence that subjects may simplify even a three-factor model by adding rather than multiplying (Shanteau & Anderson, 1969, 1972; Klitzner & Anderson, 1977).

Many integration tasks, especially in judgement decision, involve both adding and multiplying operations. Analysis of such compound models is effortless. Linear fan analysis applies to two or more factors separated by a multiplication sign; parallelism analysis applies to two or more factors separated by a plus sign. The corresponding interaction tests from the analysis of variance are applicable, providing an useful tool to diagnose the underlying integration operations (Anderson, 1981, pp. 70–72).

Cognitive algebra looks much clearer in hindsight because much of the uncertainty caused by non-parallelism may be cleared up. Parallelism has a natural interpretation as a joint support for an additive integration and a linear response. But non parallelism is ambiguous, since it could result from non-linearity in the response or from non-additivity in the integration, or from some uncertain combination of both together. In fact, non-parallelism was frequently observed. Many tasks that were hypothesised to follow an adding rule turned out to follow an averaging rule, which yields non parallelism under the condition of differential weighting, as noted by Anderson and Zalinski (1990).

Chapter 2

The averaging model

An outcome of the work on cognitive algebra is the prevalence of averaging processes. For example, the job satisfaction (Zhu & Anderson, 1991) as well as the prediction of self-efficacy (Ouédraogo & Mullet, 2001) are not added, but averaged; many tests disprove the adding rule and support the averaging rule (Anderson & Zalinski, 1990). The averaging model of information integration theory represents the subject's response to a multi stimulus situation as a weighted average.
The averaging is involved in many simplest processes, for example in serial integration. Averaging may be represented as a normalisation process. In any situation or group, the sum of the weights may be viewed as a normalising factor for several inputs (Zalinski & Anderson, 1989; Anderson, 1996).

As well as in equation 1.5, averaging essentially differs from addition because it involves a two parameters representation of each piece of information: the *scale* value, s, which represents the location of the stimulus on the dimension of response; and the *weight*, w, which represent its importance in the integrated response. This weight and value representation is crucially different from the additive step-wise integration.

Unlike the addition rule, the effect of each stimulus generally depends on the amount of prior information; also, unlike the addition rule, the same stimulus may have opposite effects, depending on whether its scale value is greater or lesser than the present response.

The averaging model represents the integrated response, r, as:

$$r = \frac{w_0 s_0 + \sum w_i s_j}{w_0 + \sum w_i} \tag{2.1}$$

that is a weighted sum of values, divided by the sum of the weights. The weighted average is taken over all operative information. The division by the sum of the absolute weights, $\sum w_i$, normalises every relative weight, $w_i / \sum w_i$, so that the sum of weights is the unity within each stimulus set. Whether this sum is the unity, $\sum w_i = 1$, there is no difference between the absolute weights and the relative ones. The independence assumption applies to the absolute weights and to the scale values. However, the relative weights, of any stimulus depend on the other stimuli in the set.

The initial state, which is represented by the parameters w_0 and s_0, or, somewhere only by the parameter c_0,

$$c_0 = \frac{w_0 s_0}{w_0 + \sum w_i}$$

which represents the prior memorial information. This is also called the prior belief or initial state, which plays a vital role in averaging theory. The initial state enables the averaging model to take account of the set-size effect in which added information of equal value can produce a more extreme response. As a consequence of the initial state, the response to a single stimulus is not in general a linear function of its scale value.

From equation 2.1, the response to a single stimulus S_i is the average of that stimulus and the initial state. It implies that r_i is not in general a linear function of s_i. Finally, although c_0 can be treated as a molar unit, it is not a unitary entity; it may be a complex field of cognitive elements, where the parameters w_0 and s_0 are the resultant of some integration operation over the internal stimulus field (Anderson, 1981, pp. 62–64).
The sum in equation 2.1 is taken over all effective stimuli. These may be discrete stimuli manipulated by the investigator in a factorial design, or discriminable attributes of a unitary stimulus, as, for example, in personal or social judgement. The sum may include the stimuli obtained from memory as well as external stimuli presented by the investigator. That is, each piece of information, although complex in structure, can be treated as a molar unit.
In spite of the fact that cognitive tasks normally involve interrelated dimensions, it seems unnecessary to assume independence among the pieces of information. In fact, unlike the linear model, the averaging model has no strong assumption about the independence of the factors. That is, the items have not to be always statistically independent.
The items of information used to support the averaging model may be the kinds of items that are likely to be mutually related in the minds of subjects, for example, personality traits, characteristics of a product, hypothetical person's attitudes and likely behaviours, intentions and results (Yamagishi & Hill, 1981, 1983).

Cognitive theory often requires the averaging model, but sometimes it may be inappropriate to determine optimal decisions, especially when the experimental design shows a lot of uninformative information. Here other models or criteria, as well as the Bayesian criteria, may be more predictive (Anderson & Zalinski, 1990).

2.1 The concept of weight

The concept of "weight" provides an interesting illustration of the inductive mode of scientific definition. The need for such a concept begins in common sense thinking, but the concept develops its proper definition and full meaning only within a theoretical definition; its meaning emerges gradually as part of the scientific process, so that the concept presents accumulating knowledge.

Intuitively, the need for a concept of weight seems clear. It seems natural and meaningful to ask, for example, whether negative information is more important than positive information. Under closer scrutiny, however, the concept of weight begins to blur into the concept of scale value. Negative information might have greater effect than positive information merely because its scale value has greater magnitude, not because of any difference in weight. That is, the concept of scale value might be enough, and a separate concept of weight might be unnecessary and unjustified.

Putting the concept of weight on a solid basis, therefore, must be distinguished from the concept of scale value at a more operational level. This is not entirely or even primarily an empirical problem, because it depends on the theoretical model: for example, adding models may not allow an identifiable distinction between weight and scale value.
For this reason, it has sometimes been argued that the concept of weight is unidentifiable and ought to be merged into scale value (Schönemann, Caf-

ferty, & Rotton, 1973). This argument seems like a faith in adding and linear models, without providing an empirical support.

At the theoretical level, each weight and scale value has a well-defined, conceptual existence within the averaging theory. The reality and definition of the concepts of weight and scale value are not merely hypothetical, because the averaging model has an empirical support. Therefore, the concept of weight has some claim to scientific validity (Anderson, 1981, sect. 1.6).

Weights are interesting especially for their dependence on diverse contextual factors. From this psychological perspective, it becomes clear that weights can not be normally required to remain constant along a given stimulus dimension.

Determinants of weight The weight parameter will be affected by many experimental manipulations. According to Anderson and Zalinski (1990), many manipulations appear to fall into four categories: *reliability*, *quantity*, *relevance*, and *salience* of information.

- *Reliability* is a probabilistic concept, referring to the subjective probability that the given information is a valid indicator. Source factors typically operate upon reliability. In the person perception, for example, source reliability can be manipulated by specifying how well or how long the source had known the person, or the number and variety of occasions on which the source had observed the person. These manipulations can be viewed as determinants of the subjective probability that the source information is correct, that is, of source reliability.

- *Quantity* of information can be defined by experimental operations, at least in simple cases. Thus, the set-size variable refers to the number of equivalent stimulus items. Analogously, the weight of an extended message will depend upon its length and aggregate content.

- *Relevance* refers to the implicational relationship between the stimulus information and the dimension of judgement. A given stimulus can be important in one judgement, unimportant in another. For example, warmth would be more relevant to judgements of sociableness rather than honesty. Relevance appears to involve similarity comparisons between the stimulus adjective and the prototype. The problem of relevance is central in the implicit personality theory.

- *Salience* refers to attentional factors. As an example, the dependence of weight on serial position can be interpreted as a salience effect, at least according to the attention hypothesis. Numerous other attentional factors, including repetition and perceptual emphasis, would also affect salience weighting.

These categories are reducible to one. Perhaps that is not possible, as there seems to be a clear distinction between reliability, which is a probabilistic concept, and relevance, which does not require any notion of probability. Nevertheless, all four categories can be subsumed under a general concept of "informativeness". That concept is immediate for the quantity of information; it seems acceptable for the other three categories on the basis that a more relevant, salient, or reliable stimulus is considered to be more informative (Anderson, 1981, pp. 271–273).

The concept of weight allow to explain the observed non parallelism. When deviation from parallelism appears, the averaging theory, that is, the averaged weight and value representation, provide a clear explanation. The averaging model can provide a simple account of many empirical effects, considering them as special cases of differential weighting. The interaction between factors, the observed cross-over curve, and the extremity effects generally reflect weight parameters that are not constant, but are related with scale value.

2.2 Differential weighting

General case In a typical application, the subject responds to a set of stimulus variables manipulated by the experimenter. For the case of three variables, A, B and C, the equation 2.1 may be written as:

$$r_{ijk} = \frac{w_0 s_0 + w_{Ai} s_{Ai} + w_{Bj} s_{Bj} + w_{Ck} s_{Ck}}{w_0 + w_{Ai} + w_{Bj} + w_{Ck}} \tag{2.2}$$

where i, j and k index the levels of the corresponding variables, for every subject, for every repeated session. The inclusion of the index subscripts in the weight parameters of the equation indicate the possibility of differential weighting.

Equal-Weight case If all levels of a factor A have the same weight, $w_{Ai} = w_A$, then the factor A is said to be equally weighted. If every factor is equally weighted, then the foregoing three-factor design may be written as:

$$r_{ijk} = \frac{w_0 s_0 + w_A s_{Ai} + w_B s_{Bj} + w_C s_{Ck}}{w_0 + w_A + w_B + w_C} \tag{2.3}$$

where w_A, w_B and w_C are the weights of the three factors; index subscripts are omitted to indicate constancy. The sum of weights in the denominator has the same value in all cells of the design and can be absorbed into the arbitrary scale unit. Now it is possible to write

$$k = w_t/[w_0 + w_A + w_B + w_C], \quad \text{where } t = A, B, C$$

Accordingly, the model has a linear form and can be written as:

$$\begin{aligned} r_{ijk} &= k \cdot s_0 + k \cdot s_{Ai} + k \cdot s_{Bj} + k \cdot s_{Ck} \\ &= k \cdot (s_0 + s_{Ai} + s_{Bj} + s_{Ck}) \end{aligned}$$

where $k \cdot s_t$ denotes gross stimulus values, as in the linear model. Essentially all results of linear models apply direct to the equal-weight case of the averaging model. The parallelism property holds, and the statistical analyses remain the same. Thus, the averaging model is easy to control when the equal-weight condition can be satisfied.

Overall Equal-Weight case A simpler model occurs when the weights of all factors are equal, that is, $w_{Ai} = w_{Bj}$ for every factor A, B, and for every level i,j. This situation simplifies the just mentioned equation to a plain linear form, where $k = 1/[w_0 + \text{number-of-factors}]$.

Differential-Weight case In some situations, evidences indicate that different levels of a given attribute may have different importance. In these situation, it is necessary to estimate a weight for each stimulus level of one or more factors. Instead of a single weight for each factor, a weight is estimated for every level of that factor. Such a model is called "differential weighting". An important complication concerns unequal weighting within an attribute dimension, since differential weighting is not recognised in standard methods of multiattribute analysis, which employs a constant weight for each attribute (Oden & Anderson, 1971).

In general, the averaging model allows each stimulus to have its own weight as well as its own scale value. The sum of the absolute weights in the denominator of equation 2.2 is therefore variable across the sets; the denominator varies from cell to cell in the design and the model becomes inherently non linear.

This non linearity makes the weight parameters identifiable, but introduces statistical problems, concerning bias, convergence, reliability, goodness of fit (Zalinski & Anderson, 1990, pp. 356–358) and it also requires a suitable methodology.

2.3 Parameters identifiability

2.3.1 Regression Approach

Regression analysis is a form of multi-attribute analysis in which an independent criterion is available. In the regression model, non comparable predictors are converted into common weight × value effects by virtue of the criterion variable. The analysis of variance model does the same, without requiring prior scaling of the predictors. But this conversion into common effects is accomplished by confounding the units of the weight and value scale.

This unit confounding, which underlies the practical utility of regression analysis, means that the weights themselves are not generally comparable. It follows that regression weights are not generally valid measures of psychological importance. So, the numerous attempts to interpret regression weights as measures of psychological importance are not generally meaningful (Anderson, 1976).

2.3.2 Self-estimation of the weights

An alternative to the regression analyses is to ask subjects to make numerical self-estimates of the importance weights (Zalinski, 1987, sect. 2.2). Multiattribute evaluations will clearly be better if the person's true weights are used rather than equal weights. If valid, such self-estimates would allow the use of very simple experimental designs.

A main question about self-estimates concerns their validity. Subjects produce direct estimates of scale value and importance, but it is needed to assess the validity of these estimates. Without a model or a valid criterion with which compare the estimates, it is difficult to establish their accuracy or validity.

A popular validation technique consists in comparing the self-estimates to the weights obtained by regression analyses. These comparisons withstand the foregoing problems with regression analyses. Furthermore, a number of studies found discrepancies between objective weights and subjective or self-estimated weights (Slovic & Lichtenstein, 1971; Slovic, Fischhoff, & Lichtenstein, 1982; Peng & Nisbett, 2000). These disagreements seem to reflect the inadequacies of the objective criteria more than the self-estimated weights (Reilly, 1996).

Functional measurement provides two practicable solution in order to assess validity of self-estimated parameters, especially when an external criterion is available.

At first, when both functional scales and self-estimates are available, the former provide a validational base for the latter (Shanteau & Anderson, 1972). Authors used linear fan analysis to test the multiplying model, Subjective Probability $\times$ Subjective Value. The model provided acceptable goodness of fit indexes, that is, validated functional scales of both subjective probability and subjective value. This kind of comparison provides a general basis to develop a self-estimation methodology.

A second way is to employ the self-estimated parameters in the model analysis. If the model passes the test of goodness of fit, that provides simultaneous support for the validity of the self-estimates. This method of joint model parameter validation is applicable especially in comparison of self-estimates with functional parameters.
However, according to Anderson (1982, sect. 6.2), there is almost a strong disadvantage to use self-estimates in model analysis. The results may not be very informative when there are substantial discrepancies from prediction. In that case, the model, the self-estimates, and the linearity of the response scale are all in doubt. Discrepancy are generally difficult to interpret, especially without the patterning constraints of a factorial design. Only a suitable design can validly provide estimations of the weight and value parameters (Anderson, 2001b).

2.3.3 Method of sub-designs

A general problem in estimation concerns identifiability and uniqueness. Some model parameters may not be estimable from the data, and others may have limited uniqueness. In the linear model applied to a factorial design, for example, weights are confounded with the scale units and so they are not generally identifiable[1].

With a suitable design, the averaging model can provide the common ratio scale estimates of the weight parameters and the common linear scale estimates of the scale parameters. On the basis of this scaling results, valid statistical comparisons can be made among both the estimated weights and

[1] For an introduction to the concept of identifiability and uniqueness, refer to Prakasa Rao (1992) or to Hendry, Lu, and Mizon (2004).

the estimated scale parameters. This allows a complete comparability of weights and values, both within and between stimulus dimensions. The outcome of these comparisons can be used as a basis for drawing conclusions about the relative importance and value of stimulus variables which may be qualitatively quite different (Zalinski, 1987, pp. 75–79).
However, a proper experimental design is necessary for a unique parameter estimation. In fact differential weighting is not recognised in standard methods of multi-attribute analysis, which employ a constant weight for each attribute.

Partial designs Uniqueness may be obtained using a family of partial designs, each of which includes only some of the variables. Estimation of w_0 and s_0 actually requires that set size or design size are varied. A chosen family of partial designs of the same size can provide uniqueness for the design variables themselves.
The method of sub-designs solves the problem of identifiability for the averaging model (Anderson, 1982, sect. 2.3.2). This method involves the joint use of sub-designs which omit one or more factors of the full factorial design. In the equal weight case of equation 2.3, with data from a regular factorial design, the averaging model becomes a linear model and the weight parameters are not usually identifiable. Complete parameter identifiability may be insured by using the factorial design. A simple factorial design, however, can be used to obtain linear scale estimates of either the weight or the scale parameters within each stimulus dimension.

Complete identifiability The general method to obtain the complete identifiability of the parameters is to adjoin selected sub-designs to a full factorial design. For example, a full three-ways, $A \times B \times C$ design may be supplemented with the three two-ways designs ($A \times B$, $A \times C$, and $B \times C$) and by the three one-way design. Similarly, a full two-ways, $A \times B$ design could be supplemented with the two one-way designs, corresponding to the two single factors.

To illustrate this method, suppose that the prior belief has zero weight equal to zero, $w_0 = 0$, and consider three attributes. These attributes are to be judged singularly and in pairs, by the same subject, within the same experimental task and session. From equation 2.2, the theoretical responses are:

$$r_1 = w_1 s_1 / w_1 = s_1 \tag{2.4a}$$
$$r_2 = w_2 s_2 / w_2 = s_2 \tag{2.4b}$$
$$r_3 = w_3 s_3 / w_3 = s_3 \tag{2.4c}$$

$$r_{12} = (w_1 s_1 + w_2 s_2)/(w_1 + w_2) \tag{2.5a}$$
$$r_{13} = (w_1 s_1 + w_3 s_3)/(w_1 + w_3) \tag{2.5b}$$
$$r_{23} = (w_2 s_2 + w_3 s_3)/(w_2 + w_3) \tag{2.5c}$$

From equations 2.4, the values of s_i are given directly by the response. These values may be substituted into equations 2.5 to solve for the weights. Since the unit of the weight scale is arbitrary, it may be fixed by setting $w_1 = 1$. Equations 2.5a, and 2.5b may then be solved for the remaining unknowns, w_2 and w_3. These values must also satisfy equation 2.5c which provides a test of goodness of fit to assess whether the model is correct and the weights are estimated validly (Wang & Yang, 1998).

Set of n stimuli Usually, in a true experimental design, there are more than one subject, many sessions, and almost repeated measures. Equation 2.4a should thus be correctly written as:

$$r_{1_{ijk}} = w_{1_{ijk}} s_{1_{ijk}} / w_{1_{ijk}} = s_{1_{ijk}} \tag{2.6}$$

where i refers to the subject, j refers to the session, and k to the repeated measure[2].

With suitable one-way repeated measures designs, it is possible to verify if each subject uses the same scale value in repeated trials. If the trial factor is no statistically significant, then this hypothesis may not be rejected. Now, value parameters of equations 2.4 may be identified. These value parameters of the averaging model can be validly estimated with a linear robust regression, in which the responses to sub-design stimulus sets are the dependent variables. This approach may prove itself useful because it introduces new information into the estimation procedure.

At this point, two-ways repeated measures designs verify whether the subject assigns the same importance to stimuli in the repeated trials. If no significant differences are provided, thus it seems correct to accept that each subject uses the same scale values and assigns the same weights to experimental stimuli between different measures. Moreover, the estimated parameters may be considered to be the weights of the model. In fact, under some conditions, it may be proved that these parameters are valid indicators for the averaging model.

---

[2] There are many articles and manuals concerning the factorial design for the Repeated Measures ANOVA: for example, the fundamental Girden (1992), or the more recent Weinfurt (2000). For a general introduction, consider Max and Onghena (1999).

Chapter 3

Criteria for model selection

Different models can explain observed data. Adding, multiplying, and averaging models could take into account the response variability. This variability, which has been neglected so far, is a serious problem in model analysis. Even if the assumptions of the model holds, observed data will not be perfectly congruent. Accordingly, it is necessary to test the goodness of fit of the analysed model; i.e., to assess whether the observed deviations from hypothesis may reasonably be attributed to the prevailing response variability or to the failure of the model assumptions.

Generally, any model can always fit the data, but it may fit poorly. The deviations from a model could reflect residual biases in the rating response, small stimulus interactions of no great importance, or the operation of some further process not included in the model. The evaluation of discrepancies between model and data is required. Unfortunately, there is no routine recipe to decide if deviations are serious.

The simplest approach is a qualitative approach, observing the factorial

graph. For example, in figure 1.2, the visual inspection may indicate the parallelism and hence an adding-type rule. In figure 1.4, the crossover refuses any addition process.
Formal statistical methods are available. Any set of data will contain noise, so the observed factorial graph will never follow perfectly the hypothesised model. Some deviations from the model will always be observed, and it is necessary to assess whether they can reasonably be attributed to prevailing noise or they reflect real disagreement with the cognitive rule.

3.1 Qualitative test

The inspection of the factorial graph provides an informal but convenient test of goodness of fit. In every design, the factorial graph can provide a visual index of prevailing response variability (Anderson, 1996, chapter 2).
As an illustration, the four solid curves in figure 1.4 exhibit near-parallelism and small point-wise fluctuations from parallelism. These fluctuations may be taken as an index of the current response variability. But, qualitative tests are much better at disproof rather than at proof. The disproof of the general adding hypothesis does not prove an averaging hypothesis. Only a formal statistical test can recognise that the crossover of the blue curve is reliable.

Anderson (2001b, sect. 21.4.1) proposes a robust qualitative test in order to distinguish between averaging and adding hypotheses: the *opposite effects test*.
The key idea is to add the same medium information to both high and low information. That should change the response in the same direction

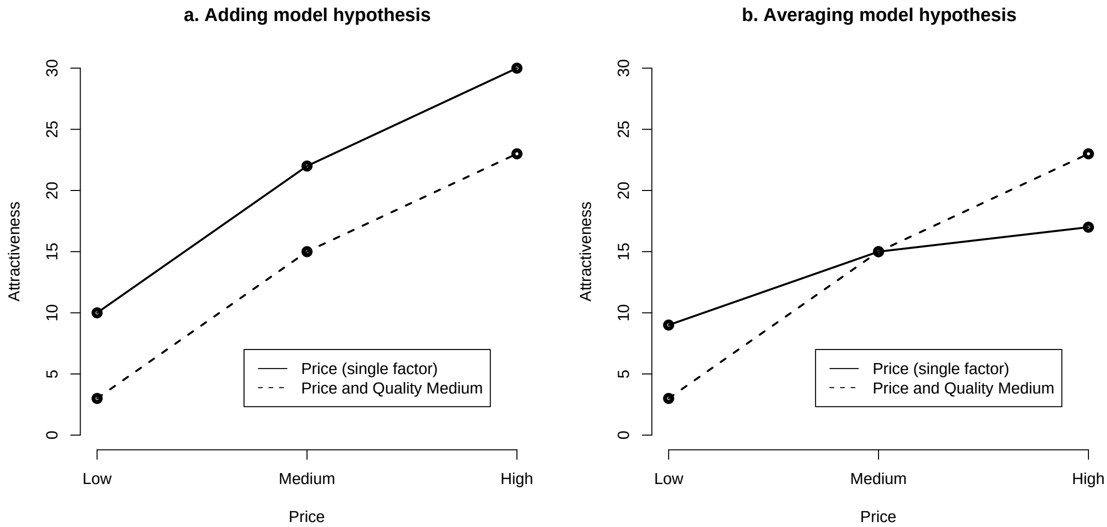


Figure 3.1: Opposite effects test.
Figure (a) supports the addition theory, while figure b. promote the averaging theory. These plots are from hypothetical data. The red dashed line represents the judgement of attractiveness as a dependent factor, and the factor "price" as a single three level independent factor. When this factor is considered with a new medium information, for example, "quality", different plots may be obtained. On the left, figure (b) shows an overall incremental effect, supporting the adding model. Instead, the crossover of dashed and solid curves in the right graph shows that the same medium information has opposite effects. It increases attractiveness whit a low price, decreases attractiveness whit a high price. This observed situation support the averaging process.

according to an adding formulation, either up or down, depending on whether the medium information is positive or negative. In contrast, an averaging formulation implies that the addition of the medium information can make the response less extreme in both cases, as shown in figure 3.1.

This test is qualitative. It depends only on the difference in direction, not in its amount. This test is robust, because only a monotone response scale is required. Furthermore, it evaluates variant forms of the adding hypothesis that may not assume an exact addition.

3.2 Equal weight case test

Under some conditions, the regression parameters, $\hat{\beta}$, which can be estimated through different approaches, are good predictors of the weights parameters in the equal weight case of the averaging model.

Suppose that prior belief has zero weight, $w_0 = 0$, and consider two attributes. The averaging equation 2.1 may be written in a linear form, where every weight correspond to an angular coefficient:

$$\begin{cases} r = \hat{\beta}_1 v_1 + \hat{\beta}_2 v_2 \\ r = \dfrac{w_1}{w_1 + w_2} v_1 + \dfrac{w_2}{w_1 + w_2} v_1 \end{cases}$$

This system requires that the condition $w_1 + w_2 \neq 0$ is satisfied.

Algebraic manipulation produces:

$$\begin{cases} \hat{\beta}_1 = \dfrac{w_1}{w_1 + w_2} \\ \hat{\beta}_2 = \dfrac{w_2}{w_1 + w_2} \end{cases} \tag{3.1}$$

Under some conditions this system of linear equations is soluble, and provides reliable weight estimators for the equal weight model obtained by the regression parameters. In fact, equation 3.1a may be written as:

$$w_2 = w_1 \frac{1 - \hat{\beta}_1}{\hat{\beta}_1}$$

Its substitution into equation 3.1b produces:

$$\begin{aligned} \hat{\beta}_2 (w_1 + (w_1 \frac{1 - \hat{\beta}_1}{\hat{\beta}_1})) &= (w_1 \frac{1 - \hat{\beta}_1}{\hat{\beta}_1}) \\ \hat{\beta}_1 \hat{\beta}_2 w_1 - w_1 (\hat{\beta}_1 - 1)(\hat{\beta}_2 - 1) &= 0 \\ w_1 (\hat{\beta}_1 + \hat{\beta}_2 - 1) &= 0 \end{aligned}$$

The stated equation might be solved setting $w_1 = 0$; but it is impossible, because $w_2 = 0$, and $w_1 + w_2 = 0$. Now, the unique solution for the system

3.1 is:

$$\hat{\beta}_1 + \hat{\beta}_2 = 1 \tag{3.2}$$

The system is still undetermined. Consider the new condition $w_1 + w_2 = k$, and $k = 1$. In fact, "in the averaging model the [relative] weights must sum to 1" (Anderson, 1981, p. 119). It is now possible to solve the system, obtaining:

$$\begin{cases} \hat{\beta}_1 = w_1 \\ \hat{\beta}_2 = w_2 \end{cases} \tag{3.3}$$

Equation 3.2 is the fundamental condition for the verification of the equal weight case of the averaging model with the regression approach. If this equation is not satisfied, then the regression parameters cannot be used to obtain the weight parameter. If this condition is satisfied, linear regression parameters are good estimators for weights in the averaging model, as expressed by the system of linear equation 3.3.

3.3 Strong inference

Although the just mentioned tests are useful, a more objective test of fit is necessary, in order to get more evidences which can support a cognitive rule and to discriminate correctly among different models.
The observed data will always show some outliers since the responses to the same stimulus vary naturally from one time to another. Thus, an error theory is needed to assess whether the non-parallelism observed in any experiment is real or merely given by the natural response variability.

Let ε denote the deviations of individual responses in every cell of the factorial design from the mean, or from the median, of that cell. Since all deviations in each cell refer to the same stimulus, their variance, σ_ε^2, is a measure of the response variability. Hence σ_ε^2 provides the general guideline to assess the goodness of fit of a model. If the observed deviations from the model are large relative to σ_ε^2, they allow to reject the model. The implementation of this last sentence requires one or more suitable indexes of deviations[1].

The ordinary analysis of variance provides a straightforward method. Parallelism is the graphical equivalent of the zero interaction term in the analysis of variance. If the parallelism theorem applies, then this statistical interaction is zero in principle and should be non significant in practice. In a similar way, if the linear fan theorem applies, then linear $\times$ linear component should be significant, as seen in section 1.3.2.

The regression and anova models are useful for outcome analysis, as well as for prediction. For process analysis, however, models based on regression have many hidden dangers (Anderson, 2001b, sect. 20.4). The decision to accept or to reject a cognitive model involves more than a statistical test of goodness of fit. A good fit may be little worthy if other models make the same prediction, involving different factors or relationship. Moreover, in model analysis, the investigator usually wishes to accept a model, that is, to accept the statistical null hypothesis that there are no significant discrepancies between the model and the observed data. But an experiment which lacks in power and has little value as evidence, may be masquerade as a success (Cohen, West, Aiken, & Cohen, 1983).

[1] Every statistical book discusses the most important fit indexes (F, R^2, the Residual Standard Error), as in the recent manual by Anderson (2001b), or by Weisberg (2005).

For example, when sample sizes are large, the significance tests are sensitive to small deviations from the null hypothesis, so that all reasonably parsimonious models may be rejected as having a statistically significant lack of fit. Standard tests are also unsuitable for comparing non-nested models and provide little guidance for choosing between models that have not been rejected. Such limitations of "classical" significance tests have stimulated interest in other approaches to model selection. One common class of such alternatives is the so-called parsimonious model selection criteria (Kuha, 2004).

3.4 The principle of parsimony as a rule for model uncertainty

Model choice is not a merely problem of goodness of fit. It concerns with decision making and with subjective and reliable criteria (Kadane & Lazar, 2004).

In modern science, there is a general agreement around a principle called "principle of parsimony". It was introduced in the Middle Ages by William of Ockham (see appendix A), and stated that a person should always opt for an explanation in terms of the lesser number possible of causes, factors, or variables. He contributed a methodological principle in explanation and theory building, especially with the formulation of a razor that bears his name, the "Ockham's razor".

In its simplest form, Ockham's Razor states that a person should make no more assumptions than what is needed. Put into in everyday language, it says: "Given two equally predictive theories, choose the simpler".

Ockham's Razor is currently considered a methodological principle, and it is often interpreted as a preference for the simplest theory that adequately

accounts for the data, because simplicity is desirable in itself (Thorburn, 1915; Ariew, 1976; Sober, 1982; Webb, 1996; Domingos, 1998).
Ockham's Razor is a basic perspective for those who follow the scientific method. It is important to note that it is like an heuristic argument that does not necessarily give correct answers; it is a loose guide in order to choose the scientific hypothesis which contains the least number of unproved assumptions.
Wrinch and Jefferys (1921) proposed to codify this theory as a rule which would automatically give an higher prior probability to laws that have fewer parameters. This approach would lead us to try at first simpler laws, moving to more complicated laws only when we find that the simple ones are not adequate to represent the data (Jefferys, 1939). That is, this approach would provide a sort of rationalised Ockham's Razor.

A frequently encountered situation is that of fitting an empirical model to data - a model that is not "true", but that will be used for prediction of the phenomenon under study. This can happen either when the "true" model is unknown or when the "true" model is too complex to be computationally useful. The selection among possible empirical models in this setting involves different considerations. The accuracy of future predictions is, of course, a major concern, but the simplicity of the model for interpretational reasons is also highly relevant. This latter factor can lead to a parsimonious Ockham's razor, which chooses the simpler model for practical reasons, not because it is true.

One of the possible Bayesian approaches to model selection is based on comparing probabilities that each of the models under consideration is the true model that generated the observed data. A similar to model uncertainty was introduced by Leamer (1978), through a Bayesian approach, to compare the models that don't fit the data equally well.

In behavioural science this approach was proposed by Raftery (1986) purely as a model selection criterion, and since then, it has been widely applied to select a single optimal model.

All results in Bayesian statistics derive directly from the definition of conditional probability, the law of total probability and the posterior distribution expressed by the Bayes' theorem (see appendix B).

The Bayesian approach is in contrast with the concept of frequency probability where the probability is derived from observed or imagined frequency distributions or proportions of populations. The difference has many implications for the methods with which statistics is practised following one model or the other, and also for the way in which conclusions are expressed.

When comparing two hypotheses and using some information, Bayesian methods suggest that one hypothesis is more probable than an other or that the expected loss associated with an hypothesis is less than the expected loss of the other. This approach strongly differs from the frequency methods in which the result is typically the rejection or non-rejection of the original hypothesis with a particular degree of confidence (type-I error). The Bayesian approach is like an extension of the ordinary logic to the degrees of belief in the range between 0 and 1.

Bayesian analysis can shed new light in the choice among models with less, more or different parameters, providing an excellent mechanisms for the selection, both for nested and for non-nested models (Hoeting, Madigan, Raftery, & Volinsky, 1999).

Comparing two models, M_1 and M_2, for an observed sample of data D, the ratio

$$\mathrm{BF}_{21} = \frac{p(D|M_2)}{p(D|M_1)} \tag{3.4}$$

is known as the Bayes factor: it is the central quantity of the Bayesian approach to model comparison. The Bayes factor is a measure of the evidence provided by the data in favor of M_2 over M_1. The ratio 3.4 can, in principle, be calculated for any pair of models for D. These need not be nested and may, in general, be completely different in form and assumptions. But, for any two models there is an infinite number of possible prior distributions and thus of Bayes factors (Kuha, 2004).

The driving idea behind this approach of model comparison is to examine the complexity of the paired models together with the goodness of how they fit the data, and to produce a measure which balances between the two.
If the observations come from a family of model whose a-priori distribution is not known exactly, the Bayesian solution consists of selecting the model which is most probable a-posteriori.

Schwarz (1978) introduced an approximation for the Bayes factor, known as the *Bayesian Information Criterion* (BIC). The generic formula of this criterion is:

$$\mathrm{BIC} = -2 \cdot loglik + log(n)k$$

Under the Gaussian error model, this becomes:

$$\mathrm{BIC} = \frac{k}{n}\ln(n) + \ln\left(\frac{\mathrm{RSS}}{n}\right) \tag{3.5}$$

where k is the number of regressors, n is the number of observations and RSS is the residual sum of squares.
In this approach, given a specified number of parameters, a likelihood ratio is

Table 3.1: Evidence for H_1

Approximate minimum t values for different grades of evidence and sample size (from Raftery, 1993, and Fischer, 2004).

		sample size			
Evidence for H_1	ΔBIC	30	100	1,000	10,000
Positive	2–6	1.84	2.15	2.63	3.03
Strong	6–10	3.07	3.20	3.59	3.90
Decisive	>10	3.66	3.82	4.11	4.38

obtained comparing one solution with all the models. An information criteria is given by penalising the models with additional parameters, following a selection criteria based on parsimony (Raftery, 1995; Burnham & Anderson, 2004). So, this method balances the complexity and the power of a model. The BIC methodology tries to find the minimal model that explains the data correctly. A model with many parameters will provide a very good fit to the data, but will have few degrees of freedom and be of limited utility. The imposition of a penalty for including too many terms in a regression model discourages the over-fitting. Thus, the preferred model is the one with the lowest value of the criterion (Burnham & Anderson, 2002).
The Bayesian criterion may be used to obtain the required value of an approximate t statistic, since it represents strong or decisive evidence. The approximate t values corresponding to different grades of evidence, and different sample size are shown in table 3.1.
Different grades of evidence can be useful to calibrate the diagnostic checks to which a model is subjected and to guide the search for a better model. In Jeffrey's view (1939), a model should not be abandoned until, in the posterior model probability sense, a better one is found.

The Bayesian approach to the model selection and accounting for the model uncertainty overcomes almost two main difficulties: the first occurs

when several nested and non-nested models may all seem reasonable given the data, but nevertheless lead to different conclusions about questions of interest. The second happens in large samples, where the P-value tests tend to reject any null hypotheses almost systematically, as opposed to the Bayesian approach.
Although a more complex model H_1 is correct, Smith and Spiegelhalter (1980, p. 216) show that the Bayesian methodology favours the simpler model H_0 only if the two models are so close that there is nothing to be lost for predictive purposes by choosing the simpler model. In this manner the BIC approach functions as a fully automatic Ockham's razor (Kass & Raftery, 1995).

3.5 R-AVERAGE function: a BIC-based algorithm for model selection

An interesting problem is concerned with the decision making process among models with differential or unequal weighting (sect. 2.2). In these models each level of each factor may have its own weight and its own value. There are several possible models which can be evaluated from simplex to complex ones.

The overall equal-weight case, $w_{Aj} = k$, for every A and j, consists in a linear model. This is a very simple model, and it may be considered as the baseline. But not necessarily does it represent the best solution, because it does not explain the crossover effect nor considers the factorial design.
A next model introduces a little non-linearity by adding one parameter to the model, that is, by differing the importance of a single weight, w_{Aj}, for any A and j. The introduction of a new parameter in the model, makes the

model itself more complex, generally providing better goodness of fit indexes, especially for the residual sum of squares. Adding another parameter, the overall fit may not be improved, although the complexity still increases and the degrees of freedom decrease.

Zalinski (1987) presented the AVERAGE program suitable to the estimation of the value × weight parameters for the unequal case of the averaging model, providing reliable estimations from the full-factorial design accompanied by all the sub-designs. This program allows to compute the weight and scale value parameters for each person individually using their responses to the stimulus configurations (i.e., information presented alone and in combination). For each participant, the program generates an absolute weight and scale value for each level of attribute, and a single weight and scale value for the initial impression. Parameter estimates are obtained by iteratively adjusting parameter values to find those that best fit the observed data by the criterion of maximum likelihood. The iterative adjustments are handled by the STEPIT function (Chandler, 1969) , a general algorithm for multivariate minimisation /maximisation. This method uses only the function values (no derivatives).

We implement the R-AVERAGE function, a program capable to provide reliable estimations for each subject as well for the sample, both from the full factorial design and from the only sub-designs.
In particular, our goal consists in the selection of the most suitable subset of the weight parameters, according to the overall goodness of fit indexes and to the complexity of the design. With the Bayesian approach, it is possible to analyse both these conditions: to test if each single weight is important

for the overall fit of the model, and to select the fundamental weights which can differ from the others. In fact, the best model is nor the linear neither the non-linear, but is the "simplex" model which better explains the data by using the smallest number of parameters.

We use a procedure described by Raftery (1995) and by Pötscher and Srinivasan (1994), whose criteria are similar to the forward stepwise regression criteria, as suggested by Kutner, Nachtsheim, Wasserman, and Neter (2004). Mainly, this procedure involves:

1. the identification of the goodness of the initial equal-weight case model with a robust parameter estimation. All the weights of this model are fixed, $w_{Aj} = k$, for every A and j. Thus, no weight is estimated;

2. the iteratively "stepping", that is a repeatedly alteration of the model at the previous step in accordance with an algorithm which consists of:

   (a) the selection of a single weight parameter to modify from the others;

   (b) the estimation of this parameter, by minimising the residual sum of squares of the non linear model whit the L-BFGS-B algorithm originally proposed by Nelder and Mead (1965) and implemented by Byrd, Lu, Nocedal, and Zhu (1995);

   (c) the direct models comparison, according to the BIC index;

   (d) the selection of this model for the next step if the evidence for the new model is at least positive ($\Delta BIC \geqslant 2$);

3. the search terminates when the stepping is no longer possible, given the stepping criteria.

The implemented R-AVERAGE function, implemented in a R program (R Development Core Team, 2005), allows:

1. to perform the Repeated Measures ANOVA, in order to test the statistical significance of the factors, before any other analysis;

2. to verify the validity of the averaging model in the equal weight case with equation 3.2;

3. to identify the model which best explains the observed data, among the adding, multiplying or the averaging model, according to the BIC index;

4. to estimate the weight and value parameters, if the design is suitable (sect. 2.3.3), both for each subject and for the sample;

5. to select the optimal subset of different weight parameters;

6. to compare the selections made by different goodness of fit indexes, i.e. the BIC index, the AIC index (Akaike, 1976), and the RSS.

7. to summarise all the results in long and short tables and to plot the estimated curves and the data.

The main functions of the program is in Appendix D.

Chapter 4

Applying model selection

In this chapter two experiments are presented. In the first, the physical knowledge is assessed with the functional measurement. The integration process of two and three variables is evaluated with the adding, multiplying and mixed models.

The second experiment presents a multi-attribute evaluation of some personal profiles concerning the factor of trust. The averaging model is assessed in the equal-weight and in the differential-weight case. The proposed methodology for the optimal model selection is implemented.

4.1 Experiment 1: Intuitive Physics

4.1.1 Common-sense physics

The intuitive or common-sense physics is concerned with the physical knowledge which operates in everyday actions, especially in motor behaviour. According to McCloskey (1983, p. 299), "everyone presumably has some sort of knowledge about motion". People have remarkably well articulated theories of motion, often with consistencies across individuals. Typically, intuitive

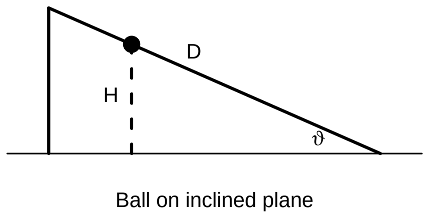


Figure 4.1: Galileo's inclined plane. Subjects predict the travel time of the ball to roll down the inclined plane.

physics regards the physical principles which govern the motions of the objects in the world (Clement, 1983; Kearney, 2002).

Intuitive physics blends perception, cognition, and action. It requires the integration of several stimulus factors, a problem to which the information integration theory may be applied (Anderson, 1983, 1997). The guiding idea is that the intuitive physics typically depends on multiple stimulus cues which are integrated to determine the overt response. For example, the prediction of the travel time of a ball in the task of the inclined plane depends on an integration of the angle of incline and the travel distance. When subjects predicts the ball's behaviour in the task of figure 4.1, the personal predictive scheme generates a functional relationship between the judgement response and the stimulus variables.

The integration function occurs within the psychological domain as a constructive process. Stimulus values do not reside in the stimulus, but are constructed by the joint process of the external stimulus field and the complexity of the internal background knowledge. This process is used in the ongoing tasks for the goals of each person.

The integration function handles the multiple determination: sensations and perceptions are the integrated resultants of multiple stimulus determinants (Anderson, 1990b, 1992). This function implies an internal representation which is concerned with the multi-dimensional relation among sensations themselves.

Much of intuitive physics may be represented in terms of schemas, which organise the information in a multiple-stimulus field and utilise background information for the operative task (Anderson, 1997).

A primary characteristic of intuitive physics is its dependence on previous experiences. People learn about the motions of objects from their earliest infancy. This basic role of a background knowledge contrasts with the *tabula rasa* approach which seeks to minimise the role of any background knowledge.

Physical judgements at all ages follow algebraic rules and these rules show a developmental trends (Wilkening & Anderson, 1982; Wilkening, Schwarz, & Rümmele, 1997; Jäger & Wilkening, 2001). When asked to judge the area of a rectangle, Area = Weight $\times$ Height, adults exhibit in the factorial graph a corresponding linear fan pattern. This pattern is the sign of a multiplying rule, which reveals a multi-dimensional concept of quantity.

In a similar task, children's data exhibit a pattern of parallelism. Anderson and Cuneo (1978a, 1978b) suggest that, although children lack of the adult conceptions of multi-dimensional quantities, they understand that a quantity judgement may be required. They seem to possess a general purpose adding rule which they apply in making judgements of a geometrical quantity (Anderson, 1980).

These developmental studies are a source of evidence for the proposition that integration rules are general-purpose functional systems and that these rules develop with increasing age. In fact, an internal representations of intuitive physics appears both in children's development and in adult learning (Schmidt & Ackermann, 1990).

The integration function, or psychological law provides the base and frame for the measurement (Anderson, 2001a). In intuitive physics, psychological measurement is even more important than in traditional psychophysics. "The stimulus integration rules [...] cannot be determined except through the operative psychological scales. In principle, functional measurement can determine the integration rules together with the psychological scales" (Anderson, 1983, p. 245).

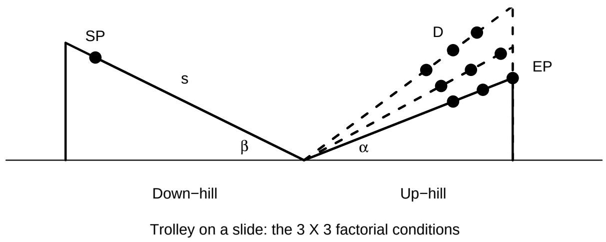
The functional measurement is capable of determining the function knowledge of intuitive physics and shows that most subjects integrate some variables, following exact addition or multiplication rules (Karpp & Anderson, 1997). The hypothesis that the stimulus integrations of intuitive physics follow algebraic rules can be tested with the parallelism theorem and the linear fan theorem (sect. 1.3).

4.1.2 Modified inclined plane

In the historical task of the inclined plane (Galileo, 1744, pp. 23–88), subjects estimate how long does it take a ball to roll down an inclined plane. Experimental factors which can be varied in the factorial design, are the distance, D, and the height, H. The distance H and D are defined in figure 4.1. With this task, Galileo proved that falling or rolling objects are accelerated independently of their mass.

Anderson (1983, 1997) presented this task for the evaluation of the underlying cognitive processes. In fact, if D and H are varied in the factorial design, the factorial graph of the physical clock times will show a linear fan pattern. In this task subjects were asked to make intuitive guesses about the travel times. These intuitive judgements were made without the benefit of any book learning. The experimental question was whether these intuitive

Figure 4.2: Task of intuitive physics.
Subjects predict the angle β (on the left) which is necessary for the ball placed in the starting point, SP, to roll down the inclined plane for a distance, s, and to go uphill, reaching the ending point, EP. Experimental factors are the angle α of the uphill slope, and the distance D, forming a 3×3 design, represented in the graph by the nine points on the right.



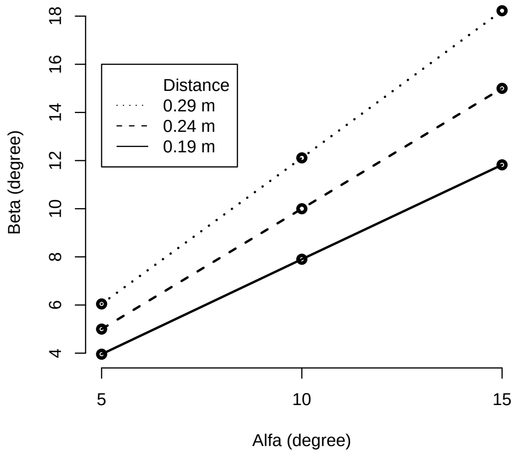
judgements will exhibit a pattern similar to the one found in nature.

One purpose of this experiment is to study the integration rules which underlie the knowledge of the motion of an object in an inclined plane. Our task is based on an experiment proposed by Bozzi (1990, pp. 296–314). Neither the travel time, nor the estimated ending point is asked to be predicted, but the angle of the downhill slope which is necessary for an object placed in a starting point, SP, to roll down the inclined plane for a constant distance, s, to go uphill on an inclined plane, for some distance, D, and to reach an ending point, EP, without passing it. The inclination of the downhill plane is determined by the angle β, and the inclination of the uphill plane by the angle α, as shown in figure 4.2.
Under ideal conditions, where, above all, there is no friction at all, the angle β, required to reach EP, is physical determined by the equation:

$$s \cdot \sin(\beta) = D \cdot \sin(\alpha) \tag{4.1}$$

If D and α are varied in the factorial design, the factorial graph of the angle β will exhibit a linear fan pattern. This is illustrate in figure 4.3.

Figure 4.3: Angle β as a function of distance and angle α.
Plot of the angle necessary for the modified inclined plane, as a function of the distance (curve parameter) and of the angle α (horizontal axis).

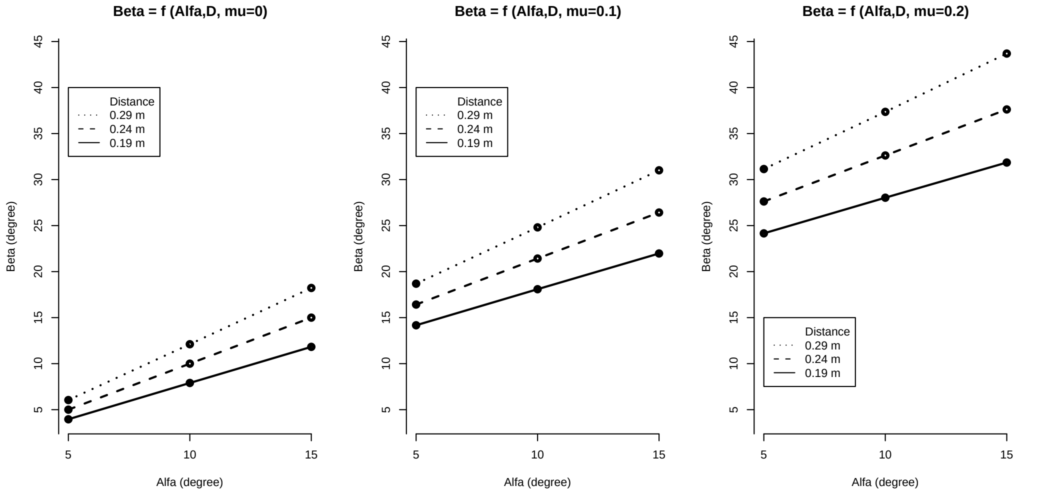


This task becomes more complex when introducing the new factor of friction. Friction is the force which opposes the relative motion of two surfaces in contact. The coefficient of friction, μ, is a scalar value which describes the ratio of the force of friction between two bodies and the force pressing them together. This coefficient is an empirical measure and cannot be found through calculations. Rougher surfaces tend to have higher values. Most dry materials in combination give the friction coefficient values from 0.1 to 0.6. It is difficult to maintain the values outside this range. A value of 0.0 means that there is no friction at all.
In order to introduce the coefficient of friction in the equation 4.1, the following equation has to be solved:

$$s \cdot (\sin(\beta) - \mu \cos(\beta)) = D \cdot (\sin(\alpha) + \mu \cos(\alpha)) \tag{4.2}$$

If μ is introduced in the factorial design, the factorial graph of the angle β will still exhibit the linear fan pattern. This is illustrate in figure 4.4.

Figure 4.4: Angle β as a function of distance, angle α, and friction.
Plots of the angle necessary for the modified inclined plane, as a function of the distance D (curve parameter), of the angle α (horizontal axis), and of the coefficient of friction μ.



4.1.3 Methods

Hypothesis Our hypotheses concern the form of the integration rule implied in the motion knowledge. The experimental question concerns the intuitive judgements, whether they exhibit a pattern similar to the one found in nature. The peculiarity of our task is the introduction of the third factor *surface* and the analysis of the changes in the integration process due to this introduction.

According to Bozzi (1990), we expect to find the multiplying or adding pattern for the integration of the factors *angle* and *distance*. With the introduction of the third factor *surface*, we look for a full multiplying pattern, $r = A \times B \times C$, where A, B, and C are the three factors, or some simpler rules, which could be yielded by some heuristics. Similarly to Singh's results (1990, 1998), the integration rules may let both the adding and the multiplying operation within the model, according to the equation $r = A + (B \times C)$, or $r = A \times (B + C)$. The integration function of these factors could use only the adding rule, showing a pattern given by $r = A + B + C$. Moreover some variables might be not considered.

Apparatus A schematic inclined plane is used. This is a rod on which a marker can be fixed at various distances. No ball rolling occurs because the aim is to study the structure of the knowledge possessed by the subjects, and the specific working rules.

Design and procedure The experiment consists of three phases, similar in purpose. The common goal is to assess the integration rules. In the first phase three *distances* (19, 24, and 29 cm) and three *slopes* (5, 10, and 15 degrees of the angle α) are factorially combined to yield the nine distance-slope stimulus configurations which are shown to the subjects. The s constant is fixed at 24 cm. Each subject judges the configurations in two successive replications. The task is described with the following instructions:

> "You can see a trolley in a slide. If you release it, the trolley will run downhill until the end of the slide. Then it will go uphill on the opposite slide. You can see a marker in this second slide. Your task consists of varying the slope of the first slide, where the trolley is situated, in order to make it reach the marker in the opposite side.
> The surface is perfectly smooth and you can imagine it as a situation in which there is no friction at all.
> Don't use formulas or other tricks you can have learnt. But, please, try to visualise the trolley going downhill and then uphill. Modify the angle of the slope and try again, until you can imagine the trolley reach the marker, but not pass it".

The experimental trials follow immediately.

In the second step we introduce the third factor *surface* with three levels: smooth, medium, rough. Thus, the subjects are asked to judge twenty-seven configuration of the 3×3×3 design. We repeat the instruction, adding these words:

"In this specific situation, you have to imagine three different kinds of surface. A smooth one, a rough one and an intermediate one, whose smothness is between the previous twos".

The third step consists of a situation where the distance D and the angle α are constant (respectively 29 cm and 15 degrees). We modify the levels of the factor *surface* from three (low, medium, high) to six, each one corresponding to a different kind of surface: ice, grass, glover, asphalt, cement, and mud, presented in a random order. Two trials are performed. The aim of this last phase is to estimate the subjective values of the friction for these different surfaces, and to look for some regularities, using the functional measurement. These regularities may give further evidences for the idea that the personal theories of motion provide adequate explanations for what we see and do.

Subjects Twenty-three subjects, eight females and fifteen males, aged from 18 to 23, mostly psychology students, take part in the experiment.

4.1.4 Results

Section I: Single subjects

Two-ways design The functional measurement shows that most of the subjects uses an algebraic rule, either addition or multiplication, to represent the joint effect of the two physical variables. Overall, the addition rule is more frequent, although it is not correct for the task.
The rule assessment is made using the repeated measures analysis of variance (ANOVA) applied separately to the data of each individual subject, using

the $\alpha = 0.05$ level of significance. As outlined in table 4.1, the functional measurement shows that only 5 out of 23 subjects used the physically correct *distance* $\times$ *slope* rule for the task. Of the remaining subjects, seven believe that only distance or slope influences the trolley motion.

We perform the repeated measures ANOVA with the *distance* $\times$ *slope* design. All subjects but three seem to correctly integrate the design factors, showing at least one factor as statistically significant (table 4.2). We exclude from the subsequent analyses the subjects number 19, 22 and 23, because none of the factors is significant, and we can not classify the rule which they use.

We analyse the goodness of fit indexes provided for the adding and multiplying models. Table 4.3 shows the BIC index and the R square values used for the model comparison. In general, no strong evidences are provided for one of the models. Subjects 1, 3, 4, 5, 6, 8, 11, 15, and 17 seem to integrate the variables with the adding rule, while subjects 2, 7, 13, 14, and 21 with the multiplying one.

Three-ways design The functional measurement shows that most of the subjects integrates the three variables following some algebraic rules. Table 4.4 shows that no subjects adopt the physics multiplying rule. Six subjects only account for two variables, without regarding the factor *distance*. Of the remaining subjects, six integrate the variables with the adding rule. The goodness of fit indexes suggest that the remaining ones adopt a mixed adding and multiplying rule. All results for each subject are reported in tables E.1–E.23.

Table 4.1: Rule assessment with functional measurement theory for the two-ways design.

Rule	N	Subjects
Distance only	3	12 13 18
Slope only	4	9 10 16 20
Distance + Slope	9	1, 3, 4, 5, 6
		8, 11, 15, 17
Distance × Slope	5	2, 7, 13, 14, 21
Unclassified	3	19 22 23

Table 4.2: Individual results.
Repeated measures ANOVA for the two-ways design *distance* × *slope*. Table reports the F-value and the p-value for each factor of the design, where significant.

Subject	Slope		Distance		Interaction	
	$F_{2,2}$	P-value	$F_{2,2}$	P-value	$F_{4,4}$	P-value
1	$1e+32$	< 0.0001	110.3	0.0008	2.6	n.s.
2	163.5	0.006	16.1	n.s.	5.4	n.s.
3	592.9	0.001	11.2	n.s.	3.4	n.s.
4	1274.3	0.001	7.8	n.s.	0.3	n.s.
5	32.3	0.030	38.2	0.0256	0.5	n.s.
6	56.7	0.017	38.9	0.0256	1.9	n.s.
7	23.2	0.041	10.7	n.s.	0.5	n.s.
8	41.6	0.023	25.8	0.037	1.6	n.s.
9	52.8	0.018	4.2	n.s.	2.9	n.s.
10	437.2	0.002	1.6	n.s.	3.1	n.s.
11	11.7	n.s.	12.0	n.s.	0.4	n.s.
12	10.9	n.s.	37.19	0.026	3.3	n.s.
13	11.9	n.s.	430.3	0.002	1.2	n.s.
14	76.4	0.013	4.9	n.s.	2.0	n.s.
15	45.9	0.021	2.3	n.s.	0.1	n.s.
16	18.2	0.05	5.6	n.s.	0.2	n.s.
17	78.1	0.012	4.3	n.s.	0.2	n.s.
18	5.1	n.s.	32.4	0.039	3.7	n.s.
19	17.3	n.s.	11.9	n.s.	4.1	n.s.
20	28.5	0.034	17.9	0.05	2.0	n.s.
21	80.9	0.012	13.4	n.s.	6.6	0.05
22	7.9	n.s.	5.1	n.s.	0.2	n.s.
23	1.6	n.s.	1.3	n.s.	0.5	n.s.

Table 4.3: Goodness of fit indexes for the model comparison.
The adding model is compared with the multiplying one, in order to explain the rule $R = f$(Distance, Slope). Table reports the BIC index and the R^2 value. The lower is the value of the BIC index the better is the model. Contrariwise, the higher is the value of R^2 the better is the model. First subject.

Subject	Distance + Slope		Distance × Slope		Evidences
	BIC	R^2	BIC	R^2	
1	53.0	0.95	67.3	0.89	Decisive
2	86.9	0.91	82.7	0.92	Strong
3	40.4	0.98	61.4	0.93	Decisive
4	54.2	0.97	73.6	0.91	Decisive
5	100.7	0.84	103.6	0.82	Positive
6	79.6	0.88	106.5	0.57	Decisive
7	83.6	0.77	81.4	0.79	Positive
8	101.4	0.82	114.8	0.63	Decisive
9	105.7	0.75	107.6	0.72	None
10	99.1	0.81	99.6	0.80	None
11	96.4	0.83	100.5	0.78	Positive
12	94.8	0.75	104.8	0.60	Decisive
13	93.6	0.80	89.9	0.83	Positive
14	95.7	0.76	91.1	0.80	Positive
15	76.8	0.89	89.2	0.79	Decisive
16	116.6	0.75	116.4	0.76	None
17	93.5	0.89	96.8	0.88	Positive
18	130.4	0.67	131.2	0.68	None
20	123.6	0.83	122.1	0.84	None
21	91.6	0.90	86.1	0.93	Positive

Table 4.4: Rule assessment with functional measurement theory for the three-ways design.

Rule	N	Subjects
Slope + Surface	6	3 7 9 11 15 20
Distance + Slope × Surface	11	4 5 6 8 10 13
		14 16 17 19 22
Distance + Slope + Surface	6	1 2 12 18 21 23
Distance × Slope × Surface	0	

Overall, the introduction of the last factor mostly changes the integration process. That is, introducing the factor *surface*, the foregoing factor *distance* is no more significant for six subjects (subject 3, 7, 9, 11, 15, and 20). The other participants correctly evaluate the differences among the levels of the factors and the three factors themselves.
We use the BIC index and the R square (R^2) values for the rule assessment. These goodness of fit indexes, reported in table 4.5, generally do not provide any strong evidence for one model. Subjects 1, 2, 12, 18, 21, and 23 seem to integrate the variables with the adding rule, while the others with an adding and multiplying one.

As discussed in 1.3.4, we also look for the integration process in each and every pair of factors. All, except subjects 5 and 14, exhibit a set of parallel curves which may be yielded by the adding rule process. But these evidences are weak for the lack of replication within subjects.

Surfaces We perform a repeated measure ANOVA to test the differences among the levels of the factor *surface*. Table 4.6 shows that all subjects except one recognise the different kinds of surface. That is, they increase the needed angle β according to the smoothness of the surfaces.

Section II: All sample

Two-ways design We analyse the whole sample in order to verify if the findings for the subjects may be applied to the sample. A repeated measures ANOVA is performed with the *distance* $\times$ *slope* $\times$ *subject* design. Results are reported in table E.24.
Figure 4.5 shows the combined effects of the factors *distance* and *slope*. On the vertical axis there are the predicted angles β. The three curves corre-

Table 4.5: Goodness of fit indexes for the model comparison.
The adding model is compared with the multiplying and the mixed ones for the factor *distance* (A), *slope* (B), and *surface* (C). Table reports the BIC index and the R^2 value. The lower is the value of the BIC index the better is the model. Contrariwise, the higher is the value of R^2 the better is the model. The subjects are reported only when all factors are significant.

Subject	Adding $A + B + C$		Multiplying $A \times B \times C$		Mixed $A + B \times C$	
	BIC	R^2	BIC	R^2	BIC	R^2
1	167.4	0.66	134.5	0.88	133.7	0.88
2	170.8	0.67	123.5	0.94	123.6	0.94
4	216.9	0.64	204.4	0.77	204.3	0.77
5	193.3	0.59	164.5	0.86	161.3	0.88
6	200.5	0.26	165.7	0.82	163.1	0.84
8	185.9	0.43	162.5	0.78	159.8	0.80
10	199.7	0.54	154.3	0.91	153.3	0.92
12	167.4	0.63	148.8	0.81	148.9	0.81
13	190.6	0.69	164.6	0.86	158.9	0.89
14	181.5	0.60	149.0	0.86	148.4	0.86
16	197.6	0.55	168.7	0.83	168.2	0.84
17	192.0	0.52	125.9	0.96	125.8	0.96
18	208.7	0.49	162.9	0.91	163.0	0.91
19	216.8	0.46	195.0	0.75	194.9	0.75
21	172.7	0.69	133.3	0.92	133.5	0.92
22	174.6	0.50	161.1	0.68	155.1	0.74
23	187.9	0.72	164.5	0.88	164.6	0.88

Table 4.6: Predicted angle β for the different kinds of surface.

Table reports the angle β predicted by each subject in the third step. Where significant we report the F and p-value of the one-way repeated measures ANOVA. Subject 3 did not complete the task.

Subject	Ice	Cement	Asphalt	Grass	Gravel	Mud	$F_{5,5}$	P
1	22.0	22.0	22.0	27.0	32.0	32.0	1.8e+30	$<$ 0.0001
2	22.0	29.0	30.0	27.0	35.5	40.0	971	$<$ 0.0001
4	16.5	19.0	19.0	53.0	53.0	77.5	76.9	0.0001
5	18.0	26.0	30.5	41.0	39.0	47.0	51.2	0.0003
6	36.0	39.5	46.0	46.5	49.0	64.5	21.5	0.0022
7	10.5	13.0	15.0	24.0	24.0	29.0	18.4	0.0031
8	29.5	35.5	31.0	36.5	39.5	42.0	10.2	0.0120
9	29.0	32.0	33.0	36.5	38.5	41.0	29.8	$<$ 0.0001
10	33.5	43.0	24.5	37.5	54.5	37.0	50.9	0.0003
11	20.5	25.5	31.5	36.0	36.0	45.5	17	0.0037
12	22.0	30.0	27.0	41.5	35.5	45.5	27.1	0.0013
13	22.0	29.5	27.5	48.0	42.0	62.5	69	0.0001
14	28.0	27.0	30.5	31.5	43.0	44.0	36.3	$<$ 0.0001
15	15.0	17.5	16.5	20.0	24.5	21.5	12.8	0.007
16	23.0	34.0	28.5	32.5	36.0	41.5	13.7	0.0061
17	26.0	27.5	41.5	49.0	35.0	27.0	43.7	$<$ 0.0001
18	28.5	35.0	36.5	48.0	66.0	62.5	209	$<$ 0.0001
19	56.0	59.5	61.5	81.0	73.5	71.5	4.91	n.s.
20	29.5	30.5	41.5	53.0	64.0	62.0	428	$<$ 0.0001
21	21.0	24.5	25.0	28.0	43.0	35.0	56.5	0.0002
22	29.0	33.0	44.5	44.5	49.0	53.0	19.8	0.0026
23	40.5	48.0	44.5	44.5	53.0	59.5	712	$<$ 0.0001

spond to the three degree of the angle α. All curves are ascending; the higher the distance D, the higher the predicted angle β ($F_{2,2} = 90.56$, $p = 0.011$). The curves are clearly separated; the higher the angle α, the higher the predicted angle β ($F_{2,2} = 204.61$, $p = 0.0048$). As expected by the just mentioned single-subjects analysis, there are mean differences among subjects ($F_{19,19} = 20.97$, $p < 0.001$).
Although not significant, the interaction between the factors *distance* and *slope* ($F_{4,4} = 5.34$, $p = 0.067$) might evidence the multiplying integration process just mentioned for some subjects.

Three-ways design Figure 4.6 shows the combined effects of the factors distance, slope, and surface. The combination of these factors in the whole sample seems to obey the general adding rule.
In each of the three panels, the curves are not at the same level on the vertical axis. In the right panel they are higher than in the left and centre panels; the higher the degree of roughness, the higher the predicted angle β ($F_{2,571} = 190.8$, $p < .0001$). None of the interactions among factors is statistical significative.

Surfaces We perform a repeated measure ANOVA in order to test the differences among the levels of the factor *surface* and the ones of the factor *subject*. According to the foregoing findings, all the factors are significant: *surface* ($F_{5,5} = 1390$, $p < 0.0001$), *subject* ($F_{21,21} = 56.69$, $p < 0.0001$), and the interaction ($F_{105,105} = 14.5$, $p < 0.0001$). These results are graphically showed in figure 4.7.

The subjects arrange the smoothness of the different surfaces in an appropriate way. The surface with low friction is the ice, followed by cement,

Figure 4.5: $R = f(\text{Distance, Slope})$.
Box-plot of the observed data from the two-ways design. It shows the effect of the factors distance, and slope on the predicted angle β. All the sample.

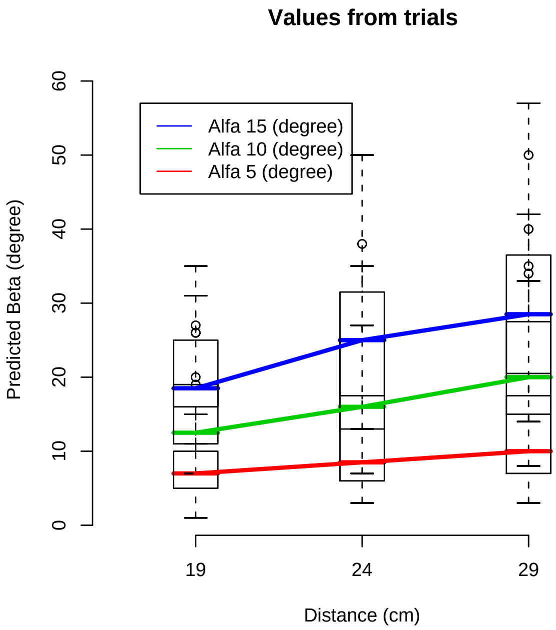


Figure 4.6: $R = f(\text{Distance, Slope, Friction})$.
Box-plot of the observed data for all the sample from the three-ways design discussed in Experiment 1. The plots show the effect of the factors surface, distance, and slope on the predicted angle β.

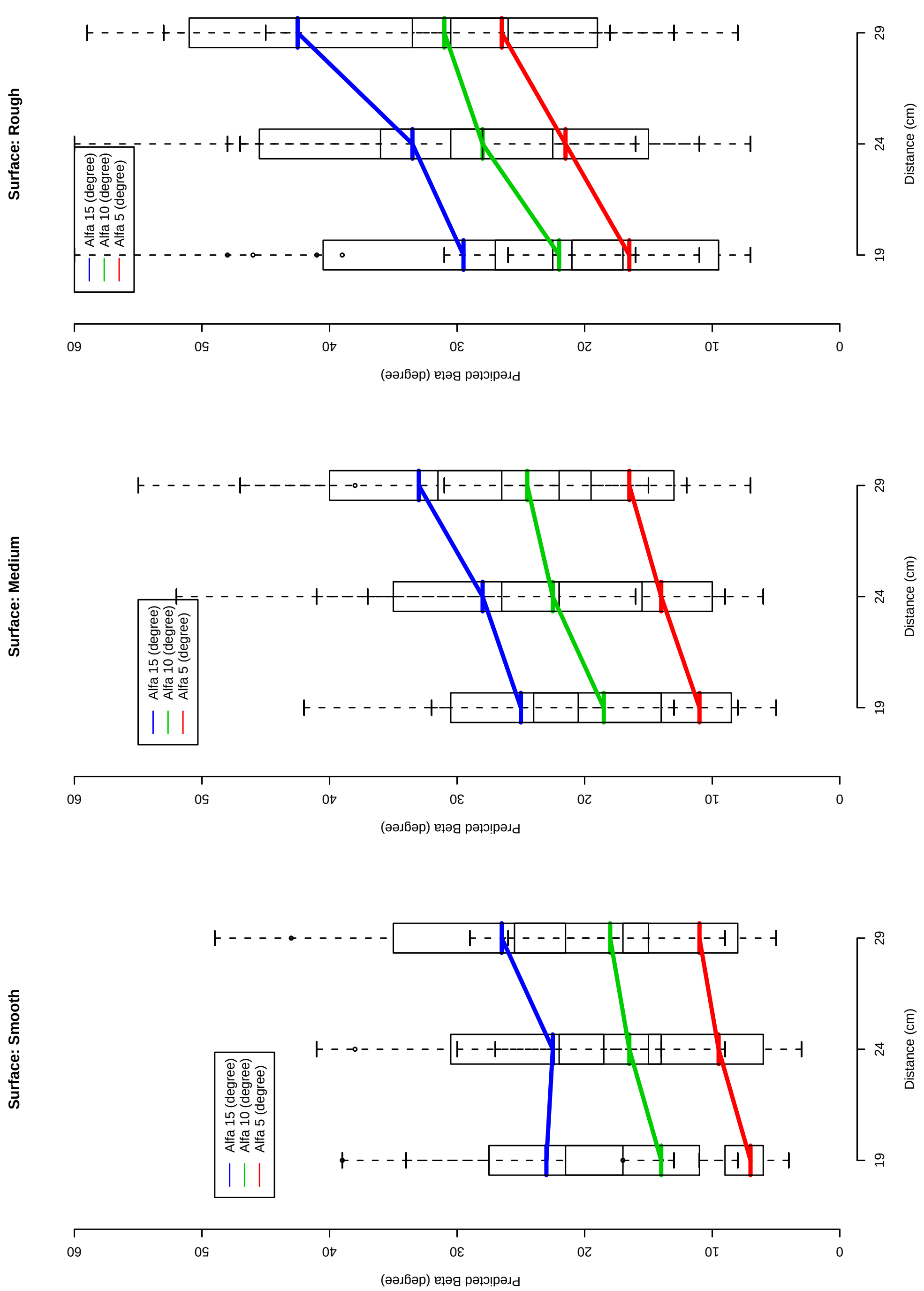
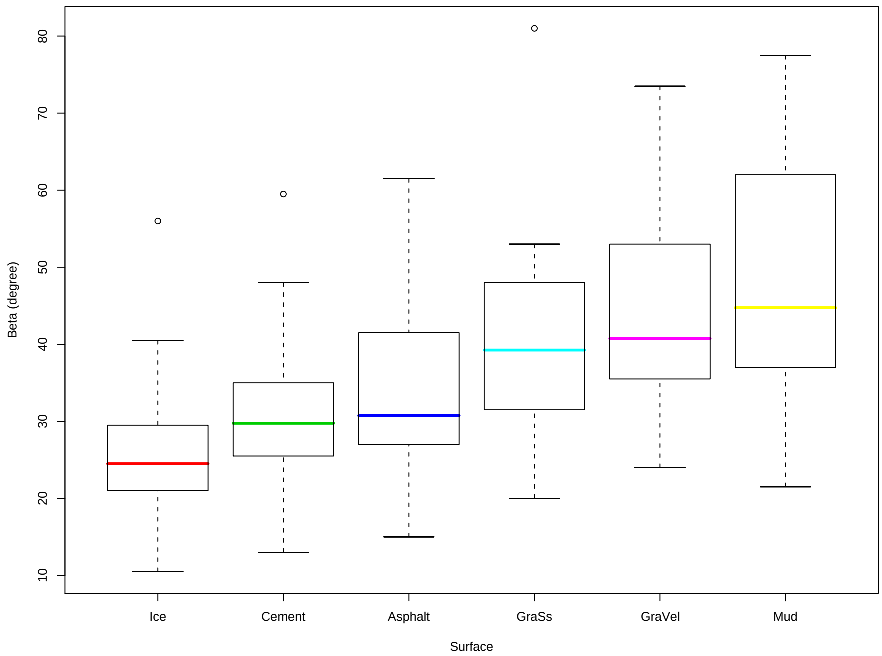


Figure 4.7: $R = f(\text{Distance})$.
Box-plot of the observed data from the third step. It shows the effect of the factors surface on the predicted angle β. All the sample.



asphalt, grass, gravel, and mud. Although the factor *surface* is significant, the subjects do not seem to differ between the level cement and asphalt (paired t-test = -0.8521, df = 21, p-value = 0.404), and between the level grass and gravel (t = -2.0797, df = 21, p = 0.0499). All the other paired t-test are significant. These findings may support the general idea that common-sense preconception in physics is not arbitrary or trivial: "every one of them was argued by pre-Newtonian intellectuals" (Kearney, 2002, p. 53).

4.1.5 Discussion

In this experiment we evaluate the function knowledge with the single subject analysis. Only if strong evidences result from this step, it is possible to look for some integration function for the whole sample. In fact, according to Karpp and Anderson (1997), the Information Integration Theory can provide a correct assessment of function knowledge especially for the single subjects and then for the sample.

In order to improve this capability, a more suitable design may be useful. For example, more replications may provide more robustness about the integration rule. Furthermore, the functional measurement may estimate the psychological values using the full-factorial design with all the sub-designs. This evaluation may refine the analysis and allow to make inferences concerning the motion perception in the population.

4.2 Experiment 2: Evaluation of trust

4.2.1 Trust models

Trust influences interactions within and among groups (Castaldo, 2002; Mayer, Davis, & Shoorman, 1995; Tyler & Kramer, 1996; Gambetta, 1988; Luna-Reyes, Cresswell, & Richardson, 2004). Trust may be defined as "an expectancy held by an individual or a group that the word, promise, verbal or written statement of another individual or group can be relied upon" (Rotter, 1967). With a more recent definition, trust is "a subjective assessment of another's influence in terms of the extent of one's perceptions about the quality and significance of another's impact over one's outcomes in a given situation, such that one's expectations [...] provide a sense of control over the potential outcomes of the situation" (Romano, 2003).
In both definitions, trust is defined in terms of expectancies or beliefs, that is, the inference about the other person's traits and intentions. Expectancies reflect the future orientation of trust. Beliefs reflect the critical role which the perceptions about the other party play in trust (Yamagishi & Yamagishi, 1994; Falcone, Pezzulo, & Castelfranchi, 2003).

Based on a trust literature review, McKnight and Chervany (1995, 2002) cluster the trusting beliefs into four attributes: *benevolence*, *honesty*, *competence*, and *predictability*.
Competence means that one believes that the other party has the ability or power to do what one needs to be done. Benevolence means that one believes that the other party cares about it and is motivated to act in one's interest. Honesty means that one believes that the other party makes good-faith agreements, tells the truth, acts ethically, and fulfils promises (Bromiley &

Cummings, 1995). Predictability means that one believes the other party's actions (good or bad) are consistent enough to be forecasted in a given situation.
Discussing a global model of trust, McKnight and Chervany (2002) suggest that these believes may be cognitive integrated in the attribute of *trustworthiness* with some weighting processes. Probably, anyone will give a high trust judgement on the people described with the attribute of high competence. But what will happen to this evaluation, whether another attribute is added? For example, low honesty, or moderate benevolence?

The following experiment is aimed at determining the integration rule which underlies the multi-attribute evaluation of trustworthiness. The main reason for distinguishing between these structures is that they have very different practical implications regarding the influence of various factors specific to each case on the propensity to trust in.
In the additive model, the impact of the different factors and the direction of the effects of the different factors (honesty, predictability, competence, benevolence) are constant, not alterable. On the contrary, in an averaging model each new factor may alter the impact of the previous factors, and the direction of the effect of this factor depends on the values of the previous factors.
For example, the presence of honesty is always a positive element even when it assumes a very weak form. By contrast, in an averaging model, this presence can be a positive or a negative element depending on the current level of honesty, which may be high as well as low.
The observation of such unexpected effects can be seen in the personality impression, where in some cases the communication of positive information

about a person can lower the general attractiveness of that person (Girard & Mullet, 1997; Girard, Mullet, & Callahan, 2002). Consequently, the difference between the two forms of the general additive model has important implications for the recommendations concerning trust management, both interpersonal and organisational (Josang, Keser, & Dimitrakos, 2005).

4.2.2 Methods

Design The full factorial design is compounded by the four trust factors (competence, benevolence, honesty, and predictability), each of them described with three levels (low, medium, high).
In order to test the exact integration process at work, we vary the number of the information factors given to the participants, as described in sect. 2.3.3. That is, we study the factorial design with six two-ways sub-designs, formed up by the combination of the four factors, and with four one-way sub-designs. This method is similar to the one discussed by Meneghelli (2004) and Zicari (2004).

Profiles The subjects are asked to judge the trustworthiness of potential people, represented by multi-attribute profiles, described as situated in similar interpersonal contexts.
The material is made up of 66 personal profiles for six interpersonal contexts: 12 single-attribute profiles and 54 two-attribute profiles. Each of the 12 single-attribute profiles mentions the single degree of honesty (low, intermediate, and high), benevolence, competence, and predictability. The 54 two-attribute profiles describe the characteristics of one person in terms of the combination of couple of attributes: honesty × benevolence, honesty × competence, honesty × predictability, benevolence × competence, benevo-

lence $\times$ predictability, competence $\times$ predictability. Each attribute has three levels. Thus, 396 profiles are presented to the subjects.
The same question appears below each profile: "How much do you believed trustworthy a person with these characteristics?" A 20 points Likert-type scale appears beneath this question. The left-hand anchor is labelled "Not at all", and the right-hand anchor, "Completely".

Procedure The profiles are presented in a random order. The participants are asked to read each interpersonal context. Then, they are asked to read each profile and to place a mark on the response scale where they believe there is the most appropriate point. The participants worked individually, at their own pace.

Subjects Four subjects are involved in this experiment. The data analysis is conducted both for each single subject and for the sample. In the investigation of single subjects, "the populations to which the drawn inferences are made up of instances of that subject and cannot validly transcend him to populations of subjects [...] Still, such single subject experiments and their logically limited conclusions can be of either practical utility or heuristic importance" (Cohen, 1988, p. 174).

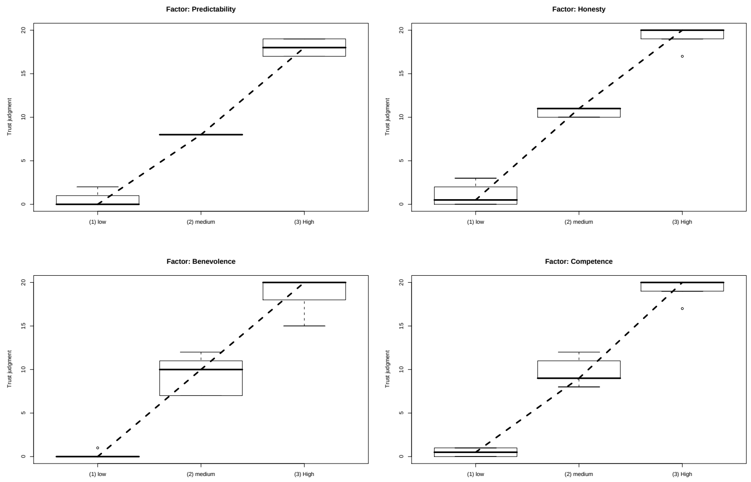
Methods of analysis The analysis consists of three phases. The first phase performs the analysis of variance on the observed data from the one-way designs. This allows to look at the significance of the personal attribute factor. We also verify the non-significance of the context factor, and we look for individual differences. Moreover, with a robust regression approach (Salibian-Barrera, 2005), we estimate the value parameters which best fit the data. We round the values to an equally distance from the medium value,

whenever this approximation does not lost power and informativeness.
In the second phases we analyse the data from the two-ways designs, in order to select the integration model, among the multiplying, adding, and averaging model.
At first, we perform the opposite effects test, associated with the equation 3.2, to verify the suitability of the averaging model. If the averaging model holds, we proceed to estimate the weights for the equal weight case with the equation 3.3. Furthermore, with the algorithm described in sect. 3.5, we look for the most suitable subset of the weight parameters for the differential weight case, according to the Bayesian goodness of fit index.
In order to improve the reliability of the selection, Burnham and Anderson (2004) suggest to verify the reliability of the BIC index, comparing it with the AIC index (Akaike, 1976), which is a different criterion for the model selection, with an underlying different philosophy: these two indexes should concord on the best model.
The third phase consists in the estimation of the overall weight and value parameters for the whole design, which includes the one-way and the two-ways designs. We expect to find the same value parameters, and the absolute weight parameters for each factor. These absolute weights could predict the integration of the factors in three-ways and four-ways factorial designs.

4.2.3 Results

The findings are presented in two main sections. In the first one, the analyses which we carry out are exemplify for the single subject 1. The second section focuses on the sample.

Figure 4.8: Values from the one-way designs.
Box-plot of the observed values for the averaging model for a single subject.

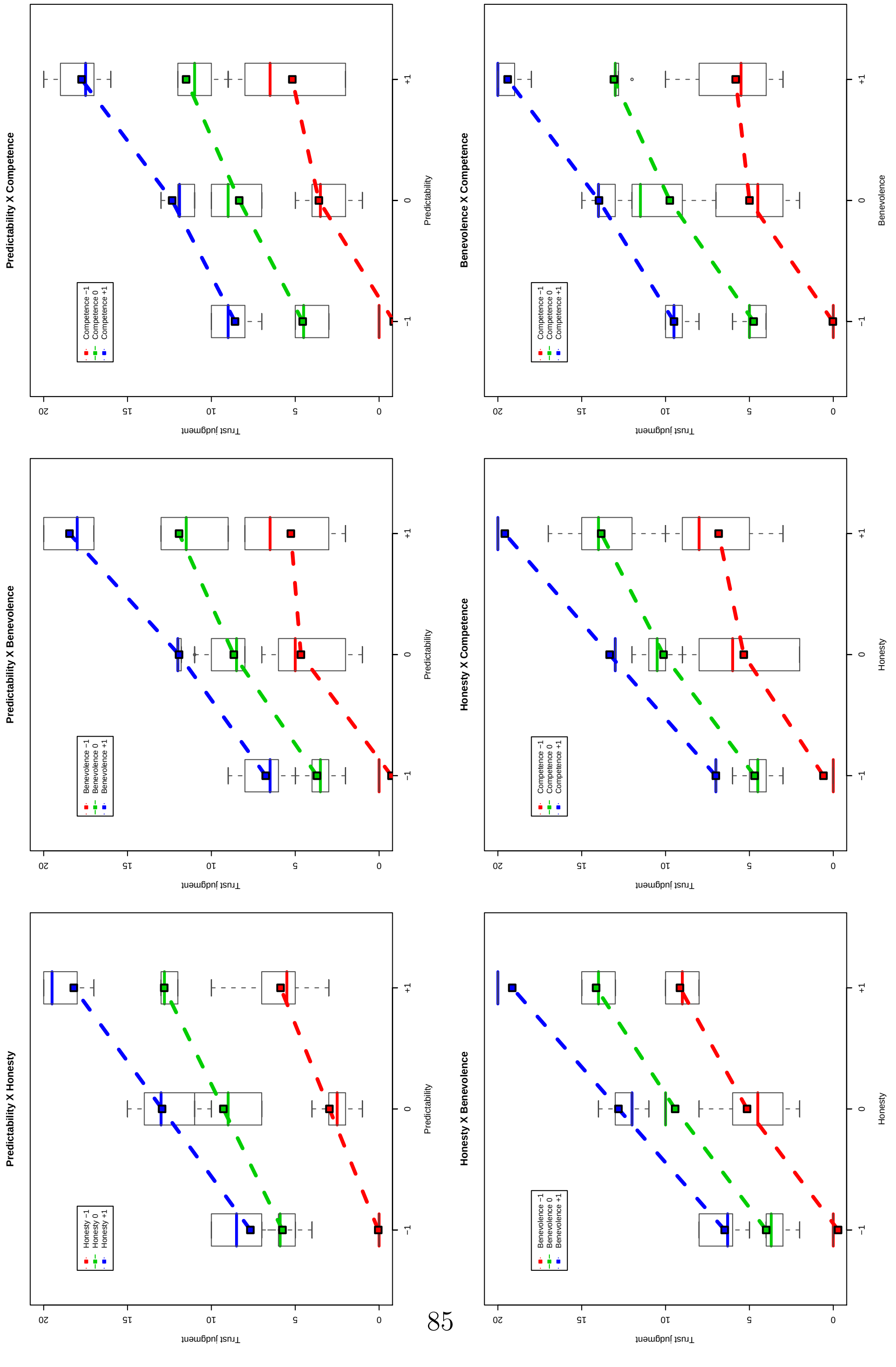


Section I: Single subject

Figure 4.8 shows the four box-and-whisker factorial plots of the observed data collected in the four one-way designs, in which we estimate the values of the factors. All the curves are clearly ascending, suggesting that the subject correctly differences among the three ordinal levels (low, medium, high). Tables E.26, and E.27 show that the different levels of the factors are always valuated with significant different values by this subject. In fact, the repeated measures ANOVA performed on the data, reports as significant all the factors: *predictability* ($F_{2,10} = 1067.3$, $p < 0.0001$), *honesty* ($F_{2,10} = 315.42$, $p < 0.0001$), *benevolence* ($F_{2,10} = 186.67$, $p < 0.0001$), and *competence* ($F_{2,10} = 409.36$, $p < 0.0001$).
Moreover, we found that the model in which the levels are rounded to the first decimal place is not significant worse than the model in which the level values are estimated by the regression. So, we retain the rounded values reported in table 4.7, as the values for the second step.

Figure 4.9 shows the factorial plots for each couple of factors. The three

Figure 4.9: Estimated parameters from the two-ways designs.
Box-plot of the observed data from the two-ways designs; dashed curves show the optimal weight and value parameters for the averaging model. Single-subject.



dashed curves represent the model which best fit the data. Tables E.28, E.29, and E.30 show that this subject generally uses the averaging rule to integrate the joint effect of the attributes, with a 90% confidence interval. This validation is obtained through the equation 3.2, which allows to verify the suitability of the averaging model through the regression parameters.
Commonly, to assess the validity of a model, the adjusted R^2 index may be usefully used. This index measures the proportion of the variation in the dependent variable accounted for by the explanatory variables. Unlike R^2, adjusted R^2 allows for the degress of freedom associated with the sums of the squares. We found that this index is always significant for all the analysed models, with very high values, > 0.95. But, in our task, this index does not allows to compare the different models. More often this index differs only at the third or more decimal place.

Table 4.8 compares the goodness of fit indexes for the different models for each sub-design. We report the residual sum of squares (RSS) and the BIC index. The lower are the indexes the better is the model. The model comparison shows that the optimal model is always obtained in the differential weight case.
In order to verify the reliability of the R-AVERAGE estimation procedure, the AVERAGE program (Zalinski & Anderson, 1986) is applied to the mean data of each sub-design. Comparing the goodness of fit indexes, the first procedure provides more reliable estimations, with ΔBIC always > 6. This difference corresponds to a strong evidence.

Table 4.9 reports the rounded estimations of the weight and value parameters for the overall design. The same procedure described in sect. 3.5

Table 4.7: Estimated values for the levels of the factors. Single subject.

	High	Medium	Low
Predictability	18.0	9.0	0.0
Honesty	20.0	10.5	1.0
Benevolence	20.0	10.0	0.0
Competence	19.5	10.0	0.5

Table 4.8: RSS and BIC index for the evaluated models for single-subject. Every sub-design is evaluated in the equally weight case, $w_A = k_A$, and in the differential weight case, $w_{Aj} = k_{Aj}$. The parameter estimations is also provided by the AVERAGE program. Factors are identified by the first uppercase letter: Predictability, Honesty, Competence, Benevolence.

	Adding		Multiplying		Averaging					
					$w_A = k_A$		$w_{Aj} = k_{Aj}$		AVERAGE	
	BIC	RSS	BIC	RSS	BIC	RSS	BIC	RSS	BIC	RSS
P×H	232	174	267	357	236	188	218	115	238	166
P×B	242	208	256	291	241	207	227	127	232	163
P×C	219	136	267	356	229	163	220	111	228	139
H×B	196	89	264	339	198	92	182	59	205	98
H×C	280	422	262	328	243	212	229	132	240	202
B×C	228	163	272	392	227	160	217	106	230	144

Table 4.9: Rounded estimations of the weight and value parameters for the overall design. Single-subject.

	Values			Weights		
	High	Med	Low	High	Med	Low
Predictability	18.0	7.5	0.0	0.5	1.5	1.0
Honesty	20.5	10.0	0.5	1.0	1.5	1.5
Benevolence	19.0	9.5	0.0	1.0	1.0	1.0
Competence	19.0	9.5	0.5	1.0	1.5	1.5

has the capability to estimate the weights from the two-ways designs and to estimate the weights and values from the overall design.
Even though we use only one-way and two-ways designs (396 profiles), without considering the three-ways sub-designs, which would require the evaluation of other 648 profiles, and the full factorial one (furthermore 648 profiles), we obtain some reliable estimations which correctly explain the observed data. With some more suitable designs, the estimated parameters would fit the data better, but the number of stimuli to be evaluated will increase quickly.
Generally, the value parameters maintain the same differences and values in all conditions. This subject does not seem to change the scale values in different conditions. The weight parameters suggest that the most important levels are the medium and low levels of the factors honesty and competence.

Figures E.1, E.2, and E.3 show similar findings for the other subjects.

Section II: All the participants

Tables E.31, E.32, E.33, and E.34 show that the different levels of the factors are always valuated with significant differences by the sample. The repeated measures ANOVA supports the significance of all the factors: *predictability* ($F_{2,10} = 99.61$, $p < 0.0001$), *honesty* ($F_{2,10} = 159.70$, $p < 0.0001$), *benevolence* ($F_{2,10} = 143.48$, $p < 0.0001$), and *competence* ($F_{2,10} = 629.37$, $p < 0.0001$). Individual differences among subjects appear for the factor *benevolence* ($F_{3,15} = 7.42$, $p < 0.003$). As for the single subjects, we keep the rounded values reported in table 4.10 for the next step.

Figure 4.10 shows the factorial plots for each couple of factors. The three dashed curves represent the model which best fits the data. Tables E.35, E.36, E.37, E.38, E.39, and E.40 show that most of the subjects use the averaging rule to represent the joint effect of the attributes.

Table 4.11 refers the estimated weights for each level of the couple of factors. We report the absolute weights setting each minimal weight to one, and multiplying the others for this value. This allows to compare the weights of the levels among themselves, in order to recognise which levels or factor are more influent in the integration process. A qualitative analysis allows to point out that, integrating the factor *predictability* with the one of three others, the subjects grant few importance to this factor against the others, with a mean ratio of 1 : 1.90, and with a maximum absolute weight equal to 9.9. Conversely, integrating the factors *competence*, *benevolence* and *honesty* in the two-ways designs, the subject give a very different value only to a bit of weight levels, with a mean ratio of 1 : 1.19, and with a maximum of 2.7 for the low level of the factor honesty integrated with the factor benevolence. These findings may suggest that the three factors are integrated with the same weights and are recognised as to be equally important.
Starting from the equal weight case, in which the levels within a factor are equal, but not the levels between the factors, the R-AVERAGE function estimates the weight parameters differing the minimal number of levels, in order to obtain the optimal set of the six parameters (three plus three weights). In the six two-ways design we obtain that the function differs from three (2+1) to six (3+3) parameters, with a mean equal to 4.5 $(2.14 + 2)$. This may indicate that some levels are more important than others, and especially that it is not always necessary to consider the full differential weight case in order to explain the averaging rule. Many time it is sufficient to consider a case in which only one or two weight levels differ.
Table 4.12 makes a comparison among the goodness of fit indexes for the

Table 4.10: Estimated values for the levels of the factors. All the sample.

	High	Medium	Low
Predictability	15.5	9.0	2.5
Honesty	19.0	9.0	1.5
Benevolence	18.0	9.5	1.0
Competence	18.0	8.5	3.0

Table 4.11: Estimated weights

The table shows the estimated weights for each level of the couple of factors, estimated in the two-ways designs by the R-AVERAGE procedure and by the AVERAGE program (in brackets). Each factor is identified by the first uppercase letter: Predictability, Honesty, Competence, Benevolence.

	1st factor			2nd factor		
	High	Med	Low	High	Med	Low
P×H	1.0 (1.0)	1.5 (2.0)	1.0 (2.0)	1.1 (2.4)	2.5 (2.1)	3.0 (2.0)
P×B	1.0 (1.0)	3.8 (2.0)	4.3 (2.0)	6.1 (2.2)	9.9 (2.0)	2.3 (2.3)
P×C	1.0 (1.0)	1.8 (1.9)	1.0 (1.9)	1.5 (2.3)	1.5 (1.9)	3.9 (1.9)
H×B	1.0 (1.0)	1.8 (1.1)	2.7 (1.1)	1.0 (1.1)	1.0 (1.1)	1.8 (1.1)
H×C	1.0 (1.9)	1.2 (2.1)	1.2 (1.0)	1.2 (1.9)	1.2 (2.1)	1.2 (1.0)
B×C	1.0 (1.4)	1.7 (2.0)	1.5 (1.0)	1.5 (1.4)	1.5 (2.1)	1.5 (1.0)

different models. This comparison shows that the optimal model is obtained in the differential weight case, although for the couple of factors *predictability* $\times$ *competence* there are not any positive evidence in order to obtain an optimal selection. In fact, the data may be explained both by the adding and the averaging model. Although the RSS of the averaging model is lower than the one of the regression, the BIC index penalises the higher number of parameters in the first model. This situation may be due to the individual differences within the subject ($F_{3,15} = 11.65$, $p < 0.0003$).

Following the procedure proposed by Falconi and Mullet (2003), we apply the AVERAGE program to the mean data of each sub-design. Generally, comparing the goodness of fit indexes, the R-AVERAGE procedure provides more reliable estimations. Only in one sub-design this program presents better estimations. But this performance is due to an adjustment of the scale values, with no regarding of the whole design.

Globally, we found found that the parameter estimated for the equal weight case accounts for the 95.55% of the variance (BIC = 1402.11). The procedure implemented by Zalinski and Anderson (1986) accounts for the 95.79% of the variance (BIC = 1411.82) and our procedure explains the most of the variance, $R^2 = 96.82\%$ (BIC = 1348.56). The general capabilities of these two algorithms will be discussed in sect. 4.3.

Table 4.13 reports the absolute parameters of the weight and value parameters for all the sample. Generally, the parameters which best fit the overall design are very similar to the ones estimated in the two-ways designs. An important result is that the value parameters maintain the same differ-

Table 4.12: RSS and BIC indexes for the evaluated models.
Every sub-design is evaluated in the equally weight case, $w_A = k_A$, and in the differential weight case, $w_{Aj} = k_{Aj}$. The parameter estimations for this last case is also provided by the AVERAGE program. Factors are identified by the first uppercase letter: Predictability, Honesty, Competence, Benevolence.

	Adding		Multiplying		Averaging					
					$w_A = k_A$		$w_{Aj} = k_{Aj}$		AVERAGE	
	BIC	RSS	BIC	RSS	BIC	RSS	BIC	RSS	BIC	RSS
P×H	1058	1534	1140	2301	1058	1538	1052	1387	1065	1477
P×B	1010	1230	1112	2028	1012	1241	983	981	1017	1180
P×C	986	1101	1083	1766	998	1163	987	1053	984	1036
H×B	1045	1449	1060	1590	1031	1355	992	1053	1014	1221
H×C	1014	1255	1030	1386	1024	1312	972	981	983	1060
B×C	1056	1521	1102	1935	1057	1531	1024	1222	1032	1265

Table 4.13: Estimations of the weight and value parameters for the overall design.

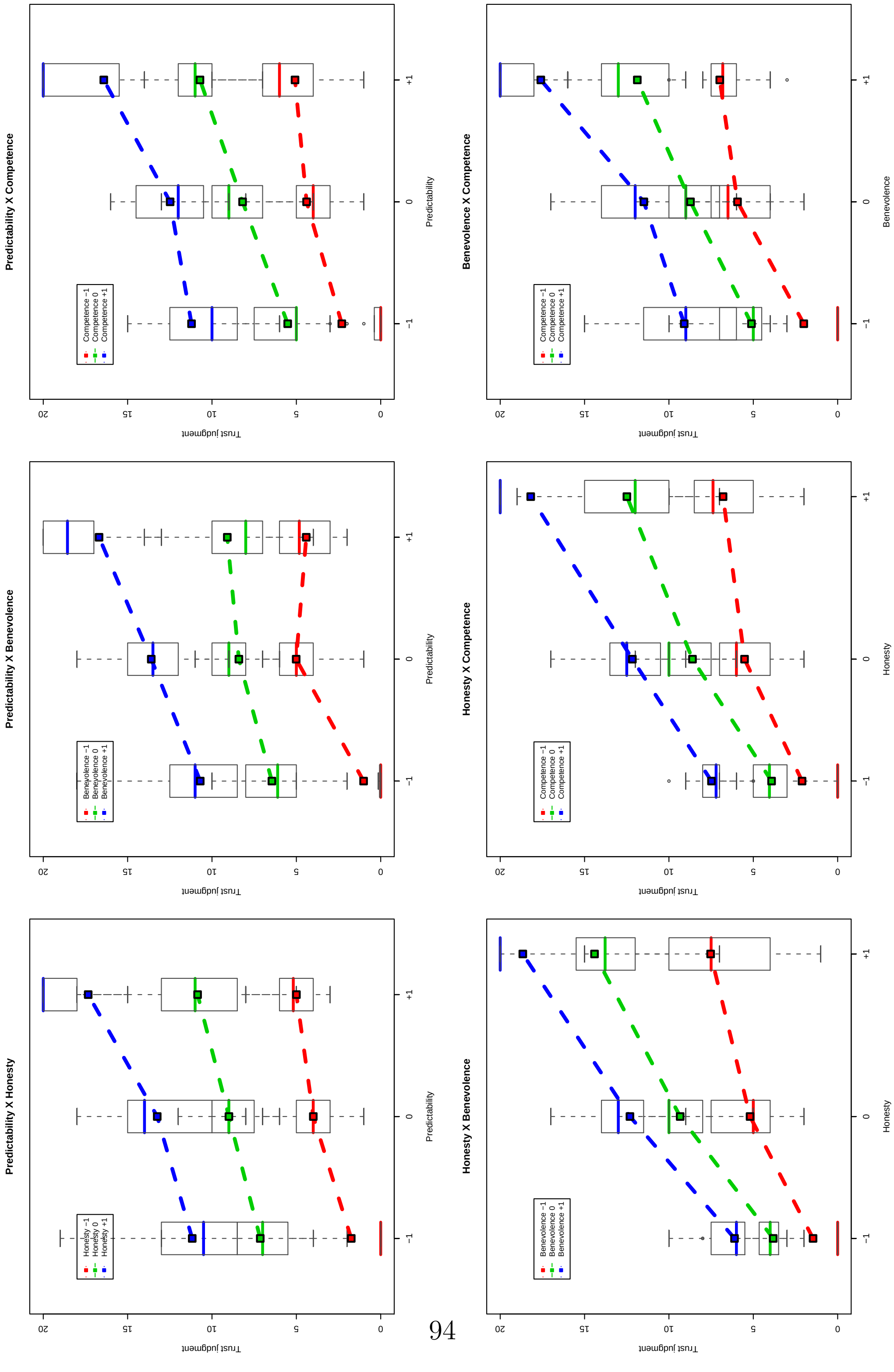
	Values			Weights		
	High	Medium	Low	High	Medium	Low
Predictability	15.3	8.5	2.0	0.7	1.6	1.1
Honesty	19.5	9.1	0.6	1.0	2.1	1.9
Benevolence	18.2	8.6	0.4	1.1	2.4	1.6
Competence	18.3	8.6	1.5	1.1	1.9	2.1

ences and values in all the conditions. Subjects do not seem to change the scale values in different conditions.

The weight parameters suggest that the most important levels are the medium levels of the factors benevolence, honesty and competence, and the low levels of the factors honesty and competence. That is, in a trust judgement, the subjects give a great negative importance to the people with low honesty or low competence, while they expect that a trustworthy person will be highly competent, honest, and benevolent.

In addition, the weights of the factor predictability are less important when the factor predictability is integrated with the other factors. This result accords with Mayer et al. (1995), who consider predictability as an economic-based sub-construct, excluding it from their trust typology.

Figure 4.10: Estimated parameters from the two-ways designs.
Box-plot of the observed data from all the sample; dashed curves show the optimal weight and value parameters for the averaging model.



4.3 Estimation procedures comparison

The last set of analyses employes a Monte Carlo simulation technique, in order to investigate the properties of the averaging model parameters estimations with two numerical procedures.

The aim of this analyses is to verify the capability of the estimated parameters to accurately define the data, and to compare the properties of the two foregoing minimisation algorithms for the averaging model: the STEPIT algorithm (Chandler, 1969) implemented in the AVERAGE program (Zalinski & Anderson, 1986), and the L-BFGS-B algorithm (Byrd, Lu, Nocedal, & Zhu, 1995), used for the R-AVERAGE function.

4.3.1 Methods

We follow a procedure similar to the one proposed by Zalinski (1987) and recently implemented by Bubna and Stewart (2000). That is, we use Monte Carlo techniques to estimate the averaging model parameters using simulated data from a standard factorial design (Manly, 1997; Wichmann & Hill, 2001a, 2001b). These simulations demonstrate that accurate estimations of the averaging model parameters can be consistently obtained from realistically simulated experiments.

Monte Carlo runs are realised by specifying the design size and true parameter values, by generating error-free data, by adding random normal error to simulate real data, and then by estimating the model parameters from these data. We analyses a 3×3 design, in which two different factors are compounded in accordance with the averaging model.

For the error-free averaging model, the data values which represent the simulated subjects' responses, range in value from 0 to 20, while the absolute weights range from 1 to 3. These ranges are similar to the findings of the previous experiment.

The random errors are obtained by generating independent normal random numbers with the algorithm proposed by Wichura (1988). We generate three sets of random error, varying the standard deviation of the errors (SD = 1, 0.1, and 1.5).

The simulations run by estimating one hundred separate sets of averaging model parameters for each set of data. This runs number is generally adequate to establish the numerical and statistical properties of the estimations (Gorin, Dodd, Fitzpatrick, & Shieh, 2005).

The parameters are estimated from the data using the two just mentioned minimisation routines and the least squares minimization criterion. The same set of bound constraints is used for each routine.

4.3.2 Results

We carry out analyses on the simulated data. In order to get a measure of the data variability, we calculate the RSS provided by the real parameters set, which is compounded by 3 degrees of values + 3 degrees of weights × 2 factors.

With the procedure implemented on the L-BGFS-B algorithm, we estimate the parameters set by fitting the averaging model to all the replication data, and we calculate the RSS. Moreover, we calculate the RSS provided by the parameters sets estimated for each single replication.

In order to compare the different procedures, we also estimate the parameters by fitting the model to the data of each replication using the AVERAGE program in the way indicated by Zalinski and Anderson (1990). We calculate the RSS both for the sets of parameters and for the set of the mean ones.

Figure 4.11 shows the results for this simulation run, in which the standard deviation of the errors is the unity (SD = 1). The three dashed curves represent the real parameters and the ones estimated by the two different procedures for the differential weight case averaging model. Similar plots are obtained for other simulation conditions, with SD = 0.1 (fig 4.12) and SD = 1.5 (fig. 4.13).
A qualitative analysis carried out on these plots does not provide very strong results. In each upper plot, the dashed curves which represents the parameters estimated with the L-BGFS-B algorithm seem to be the same as the ones which describe the real parameters. Moreover, the curves for the parameters estimated with the STEPIT minimisation algorithm are not so closer as it is for the last parameters. The two bottom plots are the two normal QQ-plots of residuals. The STEPIT algorithm seems to be more inaccurate in the parameter estimation not only for the extreme observations, the outliers, but also for the observations outside one standard deviation. But a qualitative analysis does not provide strong information in order to verify and to compare the estimation capabilities.

Tables 4.14-4.16 quantitatively compare the residuals provided by the different procedures. Generally we obtain that our procedure estimates more reliable parameters with the lowest RSS and with the highest R^2 for every simulation run. All the values parameters estimated with the L-BGFS-B algorithm are bounded within a range of 10% from the real value, and the weight are in the same range. The R^2 index is always greater than the one provided for the real data, showing the efficacy of the minimisation algorithm

which allows for the data variability. When we generate a parameters set for each replication, we obtain strong weight and value parameters. That is, in this condition we obtain the minimal RSS and the highest R^2 value. The data variability is explained by the estimated parameters more than by the real parameters. Estimating a single set of parameters for each replication, the standard distribution of the parameters provided by the R-AVERAGE function is higher than the alternative one. Thus, we encourage to perform the parameters estimations using all the data from the factorial design at the same time.
The weights, values and fit indexes are very different by using the alternative algorithm. Especially the RSS grows up twice more than data variability (RSS +84% for the first data set). The worst estimation is provided for the second run, in which the errors are the smallest (RSS $+8,884\%$). Better findings are provided for the third run, in which, due to the errors, there is a great data variability (RSS +31%).

More analyses and data generations are needed in order to document the statistical properties of the parameters estimated using our and different techniques. Different starting values and settings are to be compared. In particular, we suggest to investigate different factorial designs, varying the numbers of factors and levels. The incomplete responses to the full factorial design are also to be considered. Furthermore, some efficiency indexes have to be identified and compared.

Figure 4.11: Experiment 3: First data set (Errors SD = 1).
The three dashed curves represent the real parameters and the ones estimated for the unequal case averaging model by the alternative procedures. On the botton the QQ-plot of the residuals of the parameters estimated by the R-AVERAGE and AVERAGE functions.

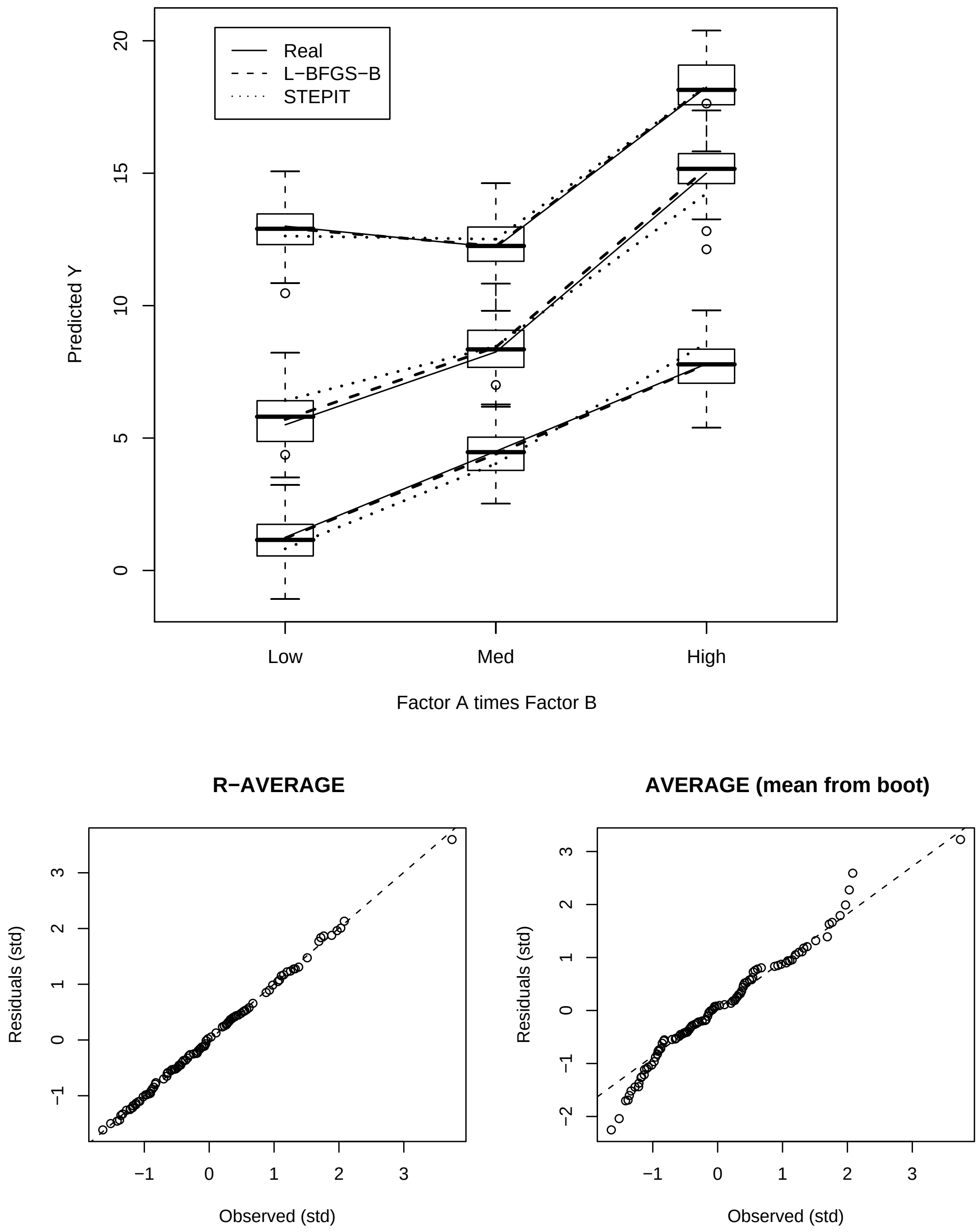


Table 4.14: Experiment 3: First data set (Errors SD = 1).

The upper table reports the estimated weight and value parameters, the R^2 index, and the residual sum of squares (RSS) estimated for all the data. The second table summaries the findings for every single replication. The mean and the standard deviation for the estimated parameters is reported.

	Factor A Values			Factor B Values			Factor A Weights			Factor B Weights			RSS	R^2
All the data	Low	Med	High	Low	Med	High	Low	Med	High	Low	Med	High		
Real parameters	2.00	8.00	18.00	1.00	9.00	18.50	1.00	3.00	2.00	3.00	1.00	2.00	1546.14	0.976
R-AVERAGE	2.10	8.17	18.08	0.94	9.03	18.39	1.00	2.98	2.17	3.10	1.00	1.84	1527.91	0.977
AVERAGE	0.16	8.36	19.72	0.97	8.61	16.89	1.00	3.09	2.98	1.51	1.00	1.03	2676.94	0.961

	Factor A Values			Factor B Values			Factor A Weights			Factor B Weights			RSS	R^2
Every replications	Low	Med	High	Low	Med	High	Low	Med	High	Low	Med	High		
R-AVERAGE														
mean	2.03	8.37	18.06	1.07	9.29	18.28	0.76	2.01	1.63	2.42	0.80	1.46	547.17	0.994
sd	0.76	0.93	0.88	0.74	0.82	0.84	1.65	2.65	2.15	3.72	1.38	2.06		
AVERAGE														
mean	0.16	8.36	19.72	0.97	8.61	16.89	0.35	1.07	1.03	1.51	1.00	1.03	2858.33	0.956
sd	0.55	1.45	0.68	0.85	1.44	1.49	0.17	0.13	0.08	0.20	0.01	0.07		

Figure 4.12: Experiment 3: Second data set (Errors SD = 0.1).

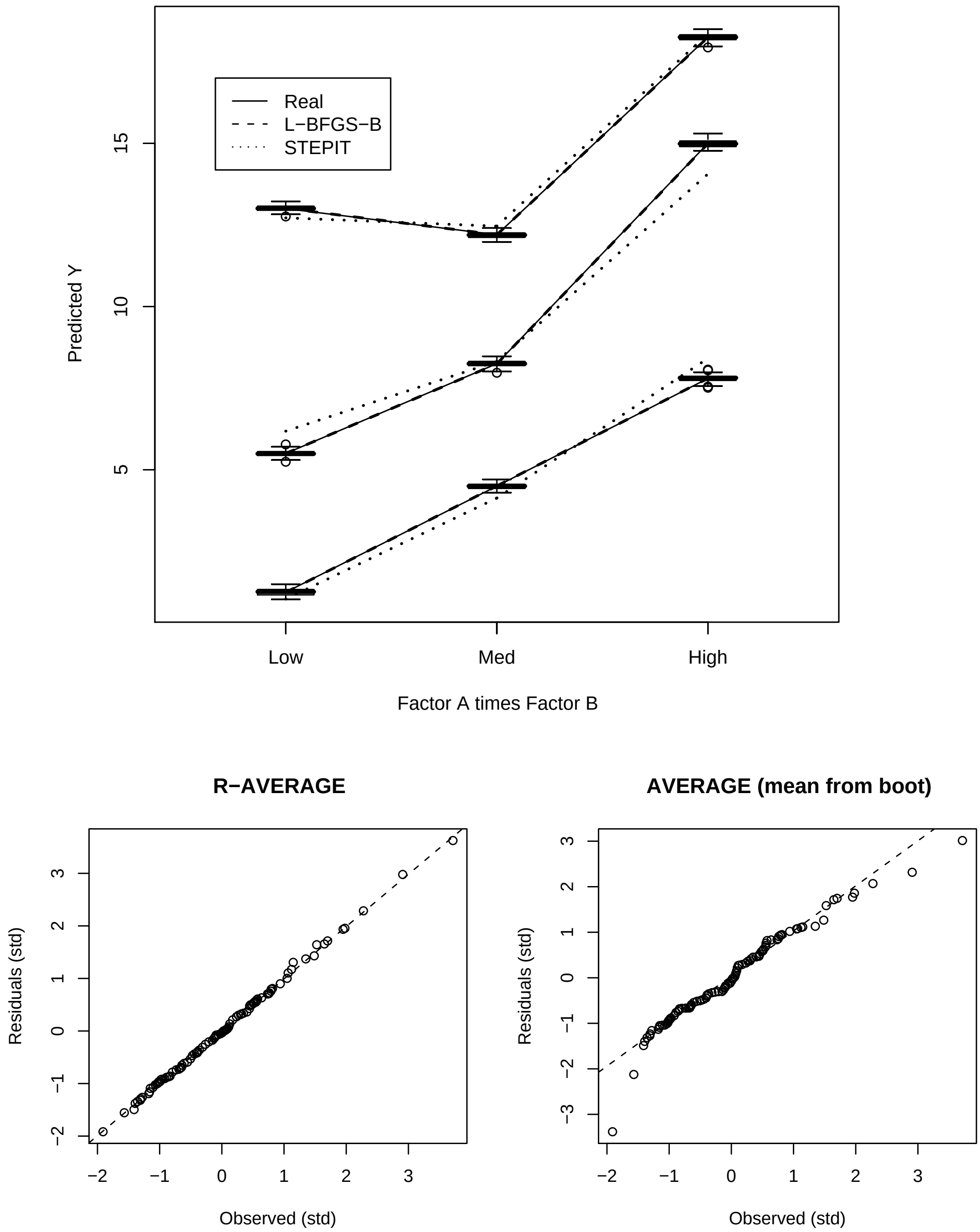


Table 4.15: Experiment 3: Second data set (Errors SD = 0.1).

The upper table reports the estimated weight and value parameters, the R^2 index, and the residual sum of squares (RSS) estimated for all the data. The second table summaries the findings for every single replication. The mean and the standard deviation for the estimated parameters is reported.

	Factor A Values			Factor B Values			Factor A Weights			Factor B Weights			RSS	R^2
All the data	Low	Med	High	Low	Med	High	Low	Med	High	Low	Med	High		
Real parameters	2.00	8.00	18.00	1.00	9.00	18.50	1.00	3.00	2.00	3.00	1.00	2.00	16.25	0.999
R-AVERAGE	2.00	7.99	10.00	0.99	8.99	18.49	1.00	3.00	2.00	2.99	1.00	2.00	16.19	0.999
AVERAGE	0.00	8.51	20.00	1.24	8.12	16.69	1.00	3.42	3.21	1.62	1.00	1.00	1457.40	0.977

	Factor A Values			Factor B Values			Factor A Weights			Factor B Weights			RSS	R^2
Every replications	Low	Med	High	Low	Med	High	Low	Med	High	Low	Med	High		
R-AVERAGE														
mean	2.00	7.99	18.00	0.99	8.99	18.49	0.64	1.92	1.28	1.91	0.64	1.29	3.79	0.999
sd	0.09	0.08	0.10	0.08	0.12	0.09	0.04	0.08	0.05	0.07	0.03	0.06		
AVERAGE														
mean	0.00	8.51	20.00	1.24	8.12	16.69	1.00	3.42	3.21	1.62	1.00	1.00	1460.80	0.976
sd	0.00	0.17	0.00	0.11	0.14	0.18	0.02	0.03	0.01	0.03	0.00	0.00		

Figure 4.13: Experiment 3: Third data set (Errors SD = 1.5).

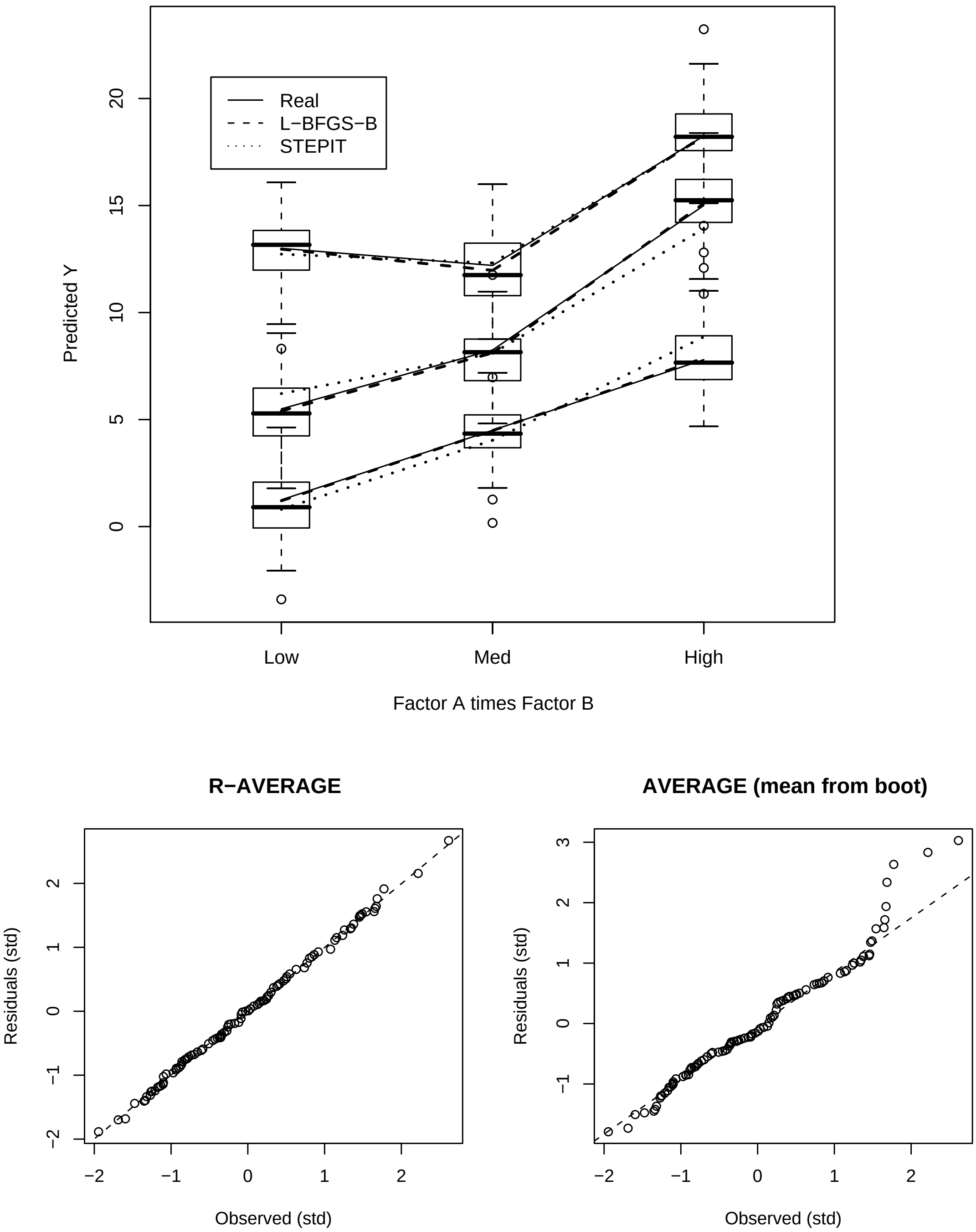


Table 4.16: Experiment 3: Third data set (Errors SD = 1.5).

The upper table reports the estimated weight and value parameters, the R^2 index, and the residual sum of squares (RSS) estimated for all the data. The second table summaries the findings for every single replication. The mean and the standard deviation for the estimated parameters is reported.

	Factor A Values			Factor B Values			Factor A Weights			Factor B Weights			RSS	R^2
All the data	Low	Med	High	Low	Med	High	Low	Med	High	Low	Med	High		
Real parameters	2.00	8.00	18.00	1.00	9.00	18.50	1.00	3.00	2.00	3.00	1.00	2.00	3845.87	0.949
R-AVERAGE	2.10	7.89	18.04	0.92	8.99	18.60	1.05	3.31	2.16	3.14	1.00	2.07	3832.029	0.950
AVERAGE	0.53	7.97	19.45	0.91	8.41	17.26	1.00	3.08	2.90	1.39	1.01	1.04	4849.71	0.934

	Factor A Values			Factor B Values			Factor A Weights			Factor B Weights			RSS	R^2
Every replications	Low	Med	High	Low	Med	High	Low	Med	High	Low	Med	High		
R-AVERAGE														
mean	1.99	8.77	18.14	1.40	9.49	18.27	0.53	1.47	1.25	1.73	0.62	1.33	2209.47	0.972
sd	0.83	1.13	0.90	0.74	1.05	0.84	0.50	0.72	0.49	0.61	0.44	0.63		
AVERAGE														
mean	0.53	7.97	19.45	0.91	8.41	17.26	1.00	3.08	2.90	1.39	1.01	1.04	5059.03	0.931
sd	1.4	2.07	1.18	1.10	1.84	1.87	0.27	0.19	0.11	0.25	0.05	0.13		

Chapter 5

Discussion

The large applicability of the cognitive algebra to model the Information Integration processes has been established for more than thirty years of research and experimentation in Psychology. Over the last years, researchers have been interested in the selection of the optimal algebraic model in order to explain the cognitive functions and the goal-oriented behaviours.

The Functional Measurement Theory allows to assess the cognitive rules especially with the support of the goodness of fit indexes. Anderson (1981, 2001b) argues that an error theory is needed to assess whether a hypothesised model accounts for the observed data and how much it explains the response variability. He suggests as statistical tools the analysis of variance (ANOVA) and the measurement of the correlation.
Moreover, the estimation of the weight and scale parameters for the purpose of quantitative and qualitative stimulus comparison has been completed in many experiments, usually with the AVERAGE program implemented by Zalinski and Anderson (1986). This program estimates the averaging parameters for a single subject, with no replications. In order to get overall

estimations, some central tendency indexes, as the mean or the median, have to be used.

The present study characterises a methodology for the improvement of the model assessment; in particular, we look for a reliable procedure for the model selection. One of our goals was to integrate the theoretical framework given by the principle of parsimony with some operative criteria, as the absolute and comparative goodness of fit indexes (R^2, F, RSS, AIC and BIC indexes).
We attempt to find the convergence of these different indexes, in order to select the optimal model which accounts for the data variability. The implemented R-AVERAGE function provides a procedure which automatically estimates and reports all these indexes and selects the best model.

Furthermore, we looked for obtaining an overall evaluation of the whole factorial design, which in general is not possible if the measurements are taken repeatedly on the same subjects. In fact, the repeated measures ANOVA, which provides a significant test for each factor and each interaction, generally lets the residual sum to zero. The zero quantity does not allow to estimate the likelihood indexes which are usually evaluated in a factorial design, and does not allow to compare different models.
But a cognitive model is plausible not only when its factors explain the observed data for a single trial, but also when this model provides some evidences for the whole, absolutely or at least towards other models. Thus, we systematically evaluate the residual from each model. Following a procedure similar to the one used by Karpp and Anderson (1997), we algebraically integrate the independent variables and than we analyse the residual between

the predicted data and the dependent variable. We evaluate the differences among the BIC index, the adjusted R^2, the RSS, and at least, the AIC index provided for the models, adjusting these indexes for the degrees of freedom of every model. This adjustment is especially required for the averaging model when comparing the equal weight case with its next ones.

A specific achievement of this research is the implemented function for the estimation of the weight and value parameters of the averaging model. Our findings suggest that the estimated parameters can account for the cognitive integration process, providing up to strong evidences for the averaging model, especially in the differential weight case.
Some preliminary findings could suggest the reliability of the estimation procedure and the suitability of the R-AVERAGE algorithm. This one provides good weight and value parameters for the averaging model under different error conditions. The estimated parameters minimise the residuals more than the ones provided by the different algorithm.
In order to estimate the model parameter, the implemented procedure does not require to constrain the data to a central tendency index; in fact, this procedure uses all the observations simultaneously in order to perform the residual minimisation.

We analyse the algebraic structure of the motion knowledge in the task of intuitive physics, and the algebraic structure of the trustworthy attribution in the task of personal judgement. Usually, we find strong evidences for cognitive algebra, especially in the single subject analysis.
In the first task, the most important finding is the general presence of an algebraic rule in motion knowledge. The functional measurement allows to

asses this rule as an adding, multiplying, or a mixed model. Furthermore, we look at the changing in the integration process due to the introduction of a new factor. The initial results suggest that, introducing a third factor, the integration function no longer considers one of the design factors, or it simplifies the task with effortless rules, which could be yielded by some heuristics. In this experiment we evaluate the intuitive physics knowledge with the single subject analysis. It is correct to look for an integration function for the sample only if the cognitive model holds for every single subject.
In the second task, supported by a suitable design, we analyse how four different attributes can be compounded in order to attribute a personal judgement of trustworthiness. We estimate both the scale values and the weight parameters for each of the trust attributes. We find the general presence of the averaging rule which integrates these factors. An interesting finding is represented by the value parameters which maintain the same values in all conditions.
Moreover, the comparison of the weight parameters suggests that some levels of the factors are more important than the others. We find that the medium and low levels of the factors benevolence, honesty and competence are more important than the high ones, and more important than the factor predictability. There are very different practical implications regarding the influence of various factors on the propensity to trust in.

The results of these experiments offer some suggestions for the researchers who are interested in estimating and comparing the adding, multiplying or averaging model and their parameters.
In order to improve the reliability of the estimations, a more suitable design may be useful. Especially, more replications may provide more robust evi-

dences about the integration rule. In fact, a way to improve the statistical properties of the model selection and parameter estimation is to obtain more data. One simple way to achieve this result is by running the same subject through multiple replications of the design configuration.

Moreover, a full-factorial design with all the sub-designs may let the functional measurement provide more reliable estimations. These improvements may refine the analysis and allow to draw inferences about populations of subjects.

Appendix A

The Ockham's razor

William of Ockham (ca. 1285-1349) was an English Franciscan friar and philosopher from a small village in Surrey, in the south-west of London. As a Franciscan, William was devoted to a life of extreme poverty. As a philosopher, he is remembered as one of the greatest logicians of all centuries.
He is considered the father of modern epistemology, because of his strongly argued position according to which only individuals exist, rather than universals, essences, or forms. He affirmed also that universals are the products of abstraction from individuals by the human mind and have no extra-mental existence (Charlesworth, 1956; Giorello, 1994). This is a critic of scholastic philosophy, whose theories grew ever more elaborate without any corresponding improvement in the predictive power.

Ockham introduced a methodological principle in the explanation and theory building, especially with the formulation of a razor that bears his name, the *Ockham's razor*. The formulation of this razor is typically phrased in Latin "entia non sunt multiplicanda praeter necessitatem", which, approximately translated, means "entities are not to be multiplied beyond necessity".

There are a variety of similar phrases in Ockham, such as: “frustra fit per plura quod potest fieri per pauciora”, “non est ponenda pluritas sine necessitate”, “quando propositio verificatur pro rebus, si duae res sufficient ad eius veritatem, superfluum est ponere tertiam”, “nulla pluralitatis est ponend nisi per rationem vel experientiam vel auctoritatem illius, qui non potuit falli nec errare, potest convinci” (Tornay, 1938; Fumagalli Beonio Brocchieri, 1996). This principle inspired numerous expressions including: the “parsimony of postulates”, and the “principle of simplicity”.

Before the 20th century it was believed that the justification for Ockham's Razor was metaphysical simplicity. It was thought that nature was, in some sense, simple and that our theories about nature should reflect that simplicity. With such a metaphysical justification came the implication that Ockham's Razor is a metaphysical principle. From the beginning of the 20th century, these views fell out of favor as scientists presented an increasingly complex view of the world. In response, philosophers turned away from metaphysical justifications for Ockham's Razor to epistemological ones including inductive, pragmatic, likelihood and probabilistic justifications. Today it is often invoked by learning theorists and statisticians as a justification of the preference of simpler models rather than more complex ones.

Scientists know from experience that Ockham's razor works, and they reflect this experience in the choice of their prior probabilities when they favour an hypothesis which is in accord with their experience. In fact, even though scientists do not usually think in terms of prior probabilities when they consider an hypothesis, they are doing actually this, that is considering simple hypotheses before the complex ones. This approach is also consistent with the tentative and the step-by-step nature of science, where an hypothesis is taken as a working hypothesis, and altered and refined as soon as new data become available (Jefferys & Berger, 1992).

"When deciding between two models which make equivalent predictions, choose the simpler one". A main problem is concerned with the equation "simplest is best". The Ockham's razor never claims to choose the "best" theory, but it only proposes simplicity as the deciding factor in the choice between two otherwise equal theories. Given more information, most of the time the more complex theory might turn out to be correct. Ockham's razor makes no explicit claims whether or not this will happen, but prompts us to use the simpler theory until we have reason to do otherwise (Murphy & Pazzani, 1994). Similarly, it is possible for two different theories to explain the data equally well, also having no relation between each other, and sharing no same elements. Some would argue that in this case Ockham's razor does not suggest any preference.

Appendix B

The Bayes' theorem

B.1 Thomas Bayes

Thomas Bayes (ca. 1702-1761) was a British mathematician and Presbyterian minister, known for having formulated a special case of Bayes' theorem, which was published posthumously (Bayes, 1763).

In the first decades of the eighteenth century, many problems concerning the probability of certain events, given specified conditions, were solved. For example, given a specified number of white and black balls in an urn, what is the probability to draw a black ball? These are sometimes called "forward probability" problems. The attention turned soon to such a problem: given that one or more balls were drawn, what can be said about the number of white and black balls in the urn? The Bayes' essay contains his solution to a similar problem, posed by Abraham de Moivre, author of The Doctrine of Chances (1733): "The probability of any event is the ratio between the value at which an expectation depending on the happening of the event ought to be computed, and the chance of the thing expected upon it's happening".

In modern utility theory, we would say that the expected utility is the probability of an event multiplying the payoff received in case of that event. Rearranging this definition for the probability, we obtain Bayes' definition. As Stigler (1983) points out, this is a subjective definition, and does not require repeated events.

B.2 The statement of Bayes' theorem

Bayes' theorem is a result of the probability theory, which relates the conditional distribution of probability to the marginal one. As a formal theorem, Bayes' theorem is valid in all interpretations of probability. Bayes' theorem relates the conditional and marginal probabilities[1] of stochastic events A and B:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{B.1}$$

In Bayes' theorem each term has a conventional name:

- $P(A)$ is the prior or marginal probability of A. It is "prior" in the sense that it does not take into account any information about B.
- $P(A|B)$ is the posterior probability of A, given B. It is "posterior" in the sense that it is derived from or entailed by the specified value of B.
- $P(B|A)$, for a specific value of B, is the likelihood function for A given B.
- $P(B)$ is the marginal probability of B, and acts as the normalizing constant.

[1] Conditional probability is the probability of an event A, given that an other event B has already occurred. Conditional probability is written $P(A|B)$, and is read "the probability of A, given B". Marginal probability means the probability of one event, regardless of the other event.

In order to derive the theorem B.1, we start from the definition of conditional probability. The probability of event A given event B is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

and the probability of event B given event A is:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Combining these two equations, we obtain:

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A)$$

Dividing both sides by $P(B)$, providing that it is non-zero, we obtain Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

More generally, for any Ai partition of the event space, the theorem can be stated as

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)} \tag{B.2}$$

The ratio $P(B|A)/P(B)$ is called the standardised likelihood, so the theorem may also be paraphrased as

$$\text{posterior} = \text{standardised likelihood} \times \text{prior}$$

In this way, Bayes' theorem can be written in terms of a likelihood ratio Λ and odds O as

$$O(A|B) = O(A) \cdot \Lambda(A|B)$$

where

$$O(A|B) = \frac{\Pr(A|B)}{\Pr(A^C|B)}$$

are the odds of A given B,

$$O(A) = \frac{\Pr(A)}{\Pr(A^C)}$$

are the odds of A by itself, and

$$\Lambda(A|B) = \frac{L(A|B)}{L(A^C|B)} = \frac{\Pr(B|A)}{\Pr(B|A^C)}$$

is the likelihood ratio.

B.3 Bayesian inference

Bayesian statisticians believe that Bayesian inference is the most suitable logical basis for the discrimination between conflicting hypotheses (Fienber, 2003). It uses an estimate of the degree of belief in a hypothesis before the advent of some evidence to give a numerical value to the degree of belief in the hypothesis after the advent of the evidence. Because it relies on subjective degrees of belief, however, it is not able to provide a completely objective account of induction.

The Bayes' theorem represents a way of incorporating new data into our understanding of the world (Stigler, 1982). Let H_i, with $i = 1, 2, \ldots, n$, be mutually exclusive and exhaustive hypotheses. Let $P(H_i|E)$ represent our personal probability that the hypothesis H_i is true, given all the relevant prior information E that is available to us. Let D represent some new piece of data that comes to our attention. Then Bayes' theorem tells us that we should update our personal probabilities according to the rule

$$P(H_i|E \circ D) = \frac{P(D|H_i \circ E)P(H_i|E)}{P(D|E)}. \tag{B.3}$$

where $P(D|H_i \circ E)$ is the probability that we would observe D, given that H_i is true and assuming the information E; and $P(H_i|E \circ D)$ is our updated personal probability that H_i is true, given both the old information E and the new information D. The denominator is the total probability of observing the data, summed over all the mutually exclusive hypotheses.

The scaling factor $P(D|H_i \circ E)P(H_i|E)/P(D|E)$ gives a measure of the impact that the observation has on the belief in the hypothesis. If it is unlikely that the observation is made unless we consider true the particular hypothesis, then this scaling factor will be large. Multiplying this scaling factor by the prior probability of the hypothesis, gives a measure of the posterior probability of the hypothesis given the observation (Jefferys & Berger, 1992).

Appendix C

A short introduction to Robust statistical procedures

Robust statistics Since 1960, many theoretical efforts have been devoted to develop statistical procedures which are resistant to small deviations from the assumptions. It is well-known that classical optimum procedures behave quite poorly under slight violations of the strict model assumptions.
Robust statistics develop themselves as an extension of the parametric statistics, taking into account that parametric models are at best only approximations to reality. Robust statistics deals with deviations from ideal models and their dangers for corresponding inference procedures. Its primary goal is the development of procedures which are still reliable and reasonably efficient under small deviations from the model, i.e. when the underlying distribution lies in a neighbourhood of the assumed model.

Main aims of robust procedures Robust statistical procedures focus on estimation, testing hypotheses and in regression models. From a data-analytic point of view, robust statistical procedures aim at:

- find the structure fitting best the majority of the data;
- identify deviating points (outliers) and substructures for further treatment;
- in unbalanced situations: identify and give a warning about highly influential data points (leverage points).

Fundamental concepts There is a great variety of approaches towards the robustness problem. Among these, the procedures based on M-estimators (and gross error sensitivity) and high breakdown point estimators (and breakdown point) play an important and complementary role. The breakdown point of an estimator is the largest fraction of the data that can be moved arbitrarily without perturbing the estimator to the boundary of the parameter space. Thus, the higher the breakdown point is, the more robust the estimator against extreme outliers grows. However, the breakdown point is not enough to assess the degree of robustness of an estimator. Instead, the gross error sensitivity gives an exact measure of the size of robustness, since it is the supremum of the influence function of an estimator, and it is a measure of the maximum effect which an observation can have on an estimator.

References This introduction is based on a tutorial on robust statistics presented by Bellio and Ventura (2005) at the International Workshop on Robust Statistics and **R**.

There are some books on robust statistics: Huber (1981) and Hampel, Ronchetti, Rousseeuw, and Stahel (1986) are the main theoretical ones; a book about practical application of robust methods with S and R functions is written by Marazzi (1993).

Appendix D

Program implementation

D.1 Experiment 1

This is the source of the program used to analyse the observed data in Experiment 1 (sect. 4.1). New classes and functions are underlined.

Listing D.1: Main function for Experiment 1.

```
source("fun.all.R")

# Read the observed dat
exp1← read.table("exp1ab.raw",header=TRUE)
exp1← as.matrix(exp1)
dati← NULL
for (i in 1:dim(exp1)[1]) {
        block← matrix(rep(exp1[i,4:6],3),3,byrow=TRUE)
        block← cbind(matrix(exp1[i,1:3],3),
                c(5,10,15),block)
        dati← rbind(dati,block) }
dati← cbind(dati[,c(1:3)],0,dati[,c(4:5)])
colnames(dati)← c("y","degree","distance",
        "texture","session","id")

# Data from the two-factor design
# Degree X Distance

trial    ← dati[dati[,"texture"]==0,]
y        ← trial[,"y"]
slope    ← factor(trial[,"degree"])
distance← factor(trial[,"distance"])
rep      ← factor(trial[,"session"])
id       ← factor(trial[,"id"])

# ANOVA Repeated measures
arm(y~distance*slope*rep*id)
# Model selection
models(trial,obj="y")
# Plot data
plot3x3(trial[,"y"],main="Values␣from␣trials")

# Data from the three factor design
# Degree X Distance X Texture

expA     ← dati[dati[,"texture"]>0,]
y        ← expA[,"y"]
```

```
slope    ← factor(expA[,"degree"])
distance← factor(expA[,"distance"])
surface ← factor(expA[,"texture"])
id      ← factor(expA[,"id"])

# ANOVA Repeated measures
arm(y~distance*slope*surface*id)
# Model selection
models(expA,obj="y")
# Plot data
y.smooth← expA[expA[,"texture"]==1,][,"y"]
y.medium← expA[expA[,"texture"]==2,][,"y"]
y.rough ← expA[expA[,"texture"]==3,][,"y"]

plot3x3(y.smooth,main="smooth",ylim=range(y))
plot3x3(y.medium,main="medium",ylim=range(y))
plot3x3(y.rough, main="rough" ,ylim=range(y))

x← values[,"y"]
roblm(y.smooth~x)
roblm(y.medium~x)
roblm(y.rough ~x)
```

D.2 Experiment 2

This is the source of the program used to analyse the observed data of each subject in Experiment 2 (sect. 4.2). Each class or function which is underlined is implemented in a separate library, available from the author.

Listing D.2: Main function for Experiment 2.

```
source("classes.R")
source("fun.all.R")

liv← 3; trial← 6; subject← 2

# Value estimation
data.single ←  readTable("one-way.raw",header=TRUE)
values      ←  matrix(NaN,4,3)

for (i in 1:length(fattore)){
        single ←  readSingle(data.single, liv, trial,
                  subject=subject, fattore=fattore[i])
        show(single)

        anova.value← valueAnova(single)
        print(anova.value)

        values← analyzeLevels(single)
        draw(values)
        print(values)
    values[i,]← extract(values)
    }

# Weights estimation
data.pair ←  readTable("two-ways.raw",header=TRUE)
pair      ←  combination(4,2)

for (i in 1:dim(pair)[1]){
        pair ←  readPair(data.pair,liv,trial,
                subject=subject,pair[i,])
        show(pair)
```

```
values.pair← paired(pair[i,],values)
anova.pesi← weightsAnova(pair,values.pair)
print(anova.pesi)

weight← analyzeWeights(pair,values.pair)
print(weight)
draw(weight)
drawOppositeTest(weight,values.pair)
}
```

Appendix E

Tables and Data plots

Here is reported the overall output of the data analysis for the first experiment, described in sect. 4.1, and for the second one, discussed in sect. 4.2.

Listing E.1: Experiment 1: Results for the 1st subject

```
# ---------------
# Subject:  1
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                  Df   Sum Sq  Mean Sq    F value    Pr(>F)
slope              2  225.778  112.889 1.1178e+32 < 2.2e-16 ***
distance           2   36.778   18.389 1.1033e+02  0.008982 **
rep                1    2.000    2.000
slope:distance     4    6.889    1.722 2.5833e+00  0.190172
slope:rep          2 2.02e-30 1.01e-30
distance:rep       2    0.333    0.167
slope:distance:rep 4    2.667    0.667
Residuals          0    0.000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.982 F(8,9): 60.625 AIC: 48.025 BIC: 56.929

[1] "Linear␣model␣A+B"
AIC index: 52.35711 BIC index: 55.9186 R2: 0.9541
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
slope      2 514.35  257.17  58.337 7.753e-09 ***
distance   2 182.00   91.00  20.643 1.722e-05 ***
surface    2 265.69  132.84  30.134 1.279e-06 ***
Residuals 19  83.76    4.41
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.9199 Sigma: 1.295

[0] Testing each couple of factors
Linear model A+B (weak)
Linear model A+C
Linear model B+C
```

```
[1] "Mixed␣model␣AxC+B"
AIC index: 129.2581 BIC index: 133.7373 R2: 0.883

[1] "Linear␣model␣A+B+C"
AIC index: 128.0060 BIC index: 134.4852 R2: 0.8794

*** Difference among models is weak
```

Listing E.2: Experiment 1: Results for the 2nd subject

```
# ---------------
# Subject:  2
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df Sum Sq Mean Sq  F value   Pr(>F)
slope               2 672.33  336.17 163.5405 0.006078 **
distance            2 162.33   81.17  16.0549 0.058634 .
rep                 1  20.06   20.06
slope:distance      4  28.33    7.08   5.4255 0.065122 .
slope:rep           2   4.11    2.06
distance:rep        2  10.11    5.06
slope:distance:rep  4   5.22    1.31
Residuals           0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.956 F(8,9): 24.579 AIC: 85.229 BIC: 94.132

[1] "Multiplying␣model␣AxB"
AIC index: 80.07035 BIC index: 82.74146 R2: 0.9215
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
slope      2 739.96  369.98 191.564 2.551e-13 ***
distance   2 137.83   68.92  35.683 3.682e-07 ***
surface    2 358.75  179.38  92.876 1.554e-10 ***
Residuals 19  36.70    1.93
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.9712 Sigma: 0.936

[0] Testing each couple of factors
Linear model A+B (weak)
Linear model A+C
Linear model B+C

[1] "Linear␣model␣A+B+C"
AIC index: 117.1326 BIC index: 123.6118 R2: 0.937

[1] "Mixed␣model␣AxC+B"
AIC index: 119.1448 BIC index: 123.624 R2: 0.937

*** Difference among models is weak
```

Listing E.3: Experiment 1: Results for the 3rd subject

```
# ---------------
# Subject:  3
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df  Sum Sq Mean Sq F value   Pr(>F)
slope               2 263.111 131.556 592.000 0.001686 **
distance            2   8.778   4.389  11.286 0.081395 .
rep                 1   0.222   0.222
slope:distance      4   1.889   0.472   3.400 0.131480
slope:rep           2   0.444   0.222
distance:rep        2   0.778   0.389
slope:distance:rep  4   0.556   0.139
Residuals           0   0.000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.993 F(8,9): 154 AIC: 31.532 BIC: 40.435

[1] "Linear␣model␣A+B"
AIC index: 39.69199 BIC index: 43.25348 R2: 0.9763
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq  F value    Pr(>F)
slope      2 302.635 151.317 179.8218 4.518e-13 ***
distance   2   5.844   2.922   3.4727 0.0518299 .
surface    2  25.729  12.865  15.2881 0.0001104 ***
Residuals 19  15.988   0.841
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.9543 Sigma: 0.001

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C (weak)

[1] "Mixed␣model␣AxC+B"
AIC index: 87.71876 BIC index: 92.19795 R2: 0.9472

[1] "Linear␣model␣A+B+C"
AIC index: 86.662 BIC index: 93.14118 R2: 0.9453

*** Difference among models is weak
```

Listing E.4: Experiment 1: Results for the 4th subject

```
# ---------------
# Subject:  4
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df    Sum Sq   Mean Sq   F value    Pr(>F)
slope               2    424.78    212.39 1274.3333 0.0007841 ***
distance            2     23.44     11.72    7.8148 0.1134454
rep                 1 6.828e-31 6.828e-31
slope:distance      4      0.22      0.06    0.3333 0.8437500
slope:rep           2      0.33      0.17
distance:rep        2      3.00      1.50
slope:distance:rep  4      0.67      0.17
Residuals           0      0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.991 F(8,9): 126.125 AIC: 44.008 BIC: 52.912

[1] "Linear␣model␣A+B"
AIC index: 54.20138 BIC index: 57.76287 R2: 0.9691
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq  F value    Pr(>F)
slope      2  561.1   280.6  156.197 1.603e-12 ***
distance   2   39.2    19.6   10.917 0.0006973 ***
surface    2 6555.3  3277.7 1824.769 < 2.2e-16 ***
Residuals 19   34.1     1.8
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.9953 Sigma: 1.268

[0] Testing each couple of factors
Linear model A+B (weak)
Linear model A+C
Linear model B+C

[1] "Mixed␣model␣AxC+B"
AIC index: 224.6418 BIC index: 229.1210 R2: 0.4222

[1] "Linear␣model␣A+B+C"
AIC index: 222.6430 BIC index: 229.1221 R2: 0.4222

*** Difference among models is weak
```

Listing E.5: Experiment 1: Results for the 5th subject

```
# ---------------
# Subject:  5
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df  Sum Sq Mean Sq F value  Pr(>F)
slope               2 884.78  442.39 32.2389 0.03009 *
distance            2 131.44   65.72 38.1613 0.02554 *
rep                 1  93.39   93.39
slope:distance      4  11.22    2.81  0.4833 0.75072
slope:rep           2  27.44   13.72
distance:rep        2   3.44    1.72
slope:distance:rep  4  23.22    5.81
Residuals           0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.874 F(8,9): 7.836 AIC: 108.944 BIC: 117.848

[1] "Multiplying␣model␣AxB"
AIC index: 100.9484 BIC index: 103.6195 R2: 0.8182

[1] "Linear␣model␣A+B"
AIC index: 100.0617 BIC index: 103.6232 R2: 0.8398

*** Difference between models is weak

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
slope      2 1106.66  553.33 33.5077 5.884e-07 ***
distance   2  279.67  139.83  8.4678  0.002348 **
surface    2  935.29  467.65 28.3191 1.996e-06 ***
Residuals 19  313.76   16.51
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.8809 Sigma: 3.167

[0] Testing each couple of factors
Multiplying model AxB (weak)
Linear model A+C
Linear model B+C

[1] "Mixed␣model␣AxC+B"
AIC index: 156.7902 BIC index: 161.2693 R2: 0.8794

[1] "Mixed␣model␣C+AxB"
AIC index: 154.6256 BIC index: 163.1047 R2: 0.8695

*** Difference among models is weak
```

Listing E.6: Experiment 1: Results for the 6th subject

```
# ---------------
# Subject:  6
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df Sum Sq Mean Sq F value  Pr(>F)
slope               2 170.11   85.06 56.7037 0.01733 *
distance            2 363.11  181.56 38.9048 0.02506 *
rep                 1   8.00    8.00
slope:distance      4  20.22    5.06  1.8958 0.27539
slope:rep           2   3.00    1.50
distance:rep        2   9.33    4.67
slope:distance:rep  4  10.67    2.67
Residuals           0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.947 F(8,9): 20.085 AIC: 80.867 BIC: 89.771

[1] "Linear␣model␣A+B"
AIC index: 78.94625 BIC index: 82.50774 R2: 0.8846
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
slope      2 770.75  385.38 23.0118 8.394e-06 ***
distance   2 729.21  364.60 21.7713 1.215e-05 ***
surface    2 331.92  165.96  9.9098  0.001128 **
Residuals 19 318.19   16.75
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.852 Sigma: 3.637

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C

[1] "Mixed␣model␣AxC+B"
AIC index: 158.6689 BIC index: 163.1481 R2: 0.835
```

Listing E.7: Experiment 1: Results for the 7th subject

```
# ---------------
# Subject:  7
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df  Sum Sq Mean Sq F value  Pr(>F)
slope               2 196.000  98.000 23.2105 0.04130 *
distance            2  33.333  16.667 10.7143 0.08537 .
rep                 1   1.389   1.389
slope:distance      4  16.667   4.167  0.4967 0.74271
slope:rep           2   8.444   4.222
distance:rep        2   3.111   1.556
slope:distance:rep  4  33.556   8.389
Residuals           0   0.000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.841 F(8,9): 5.952 AIC: 88.165 BIC: 97.069

[1] "Multiplying␣model␣AxB"
AIC index: 78.75636 BIC index: 81.42748 R2: 0.7953

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
slope      2 290.881 145.440 37.0449 2.777e-07 ***
distance   2  15.542   7.771  1.9794  0.165634
surface    2  47.466  23.733  6.0450  0.009296 **
Residuals 19  74.595   3.926
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.8259 Sigma: 1.481

[0] Testing each couple of factors
Linear model A+B (weak)
Linear model A+C
Linear model B+C (weak)

[1] "Mixed␣model␣AxC+B"
AIC index: 115.8922 BIC index: 120.3714 R2: 0.8221

[1] "Linear␣model␣A+B+C"
AIC index: 114.0388 BIC index: 120.5180 R2: 0.8255

*** Difference among models is weak
```

Listing E.8: Experiment 1: Results for the 8th subject

```
# ---------------
# Subject:  8
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                  Df Sum Sq Mean Sq F value  Pr(>F)
slope              2 388.00  194.00 41.5714 0.02349 *
distance           2 490.33  245.17 25.8070 0.03730 *
rep                1  98.00   98.00
slope:distance     4  16.67    4.17  1.5625 0.33801
slope:rep          2   9.33    4.67
distance:rep       2  19.00    9.50
slope:distance:rep 4  10.67    2.67
Residuals          0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.867 F(8,9): 7.349 AIC: 107.615 BIC: 116.518

[1] "Linear␣model␣A+B"
AIC index: 101.3618 BIC index: 104.9233 R2: 0.8171
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
slope      2 482.93  241.47 13.1773 0.0002572 ***
distance   2 246.43  123.22  6.7242 0.0061920 **
surface    2 552.30  276.15 15.0701 0.0001201 ***
Residuals 19 348.16   18.32
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.7864 Sigma: 3.115

[0] Testing each couple of factors
Linear model A+B (weak)
Linear model A+C
Linear model B+C

[1] "Mixed␣model␣AxC+B"
AIC index: 155.3328 BIC index: 159.8120 R2: 0.8034
```

Listing E.9: Experiment 1: Results for the 9th subject

```
# ---------------
# Subject:  9
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df Sum Sq Mean Sq F value  Pr(>F)
slope               2 463.44  231.72 52.7975 0.01859 *
distance            2 236.44  118.22  4.1890 0.19272
rep                 1 107.56  107.56
slope:distance      4  21.22    5.31  2.9385 0.16067
slope:rep           2   8.78    4.39
distance:rep        2  56.44   28.22
slope:distance:rep  4   7.22    1.81
Residuals           0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.8 F(8,9): 4.507 AIC: 112.528 BIC: 121.432

[1] "Multiplying␣model␣AxB"
AIC index: 104.9315 BIC index: 107.6026 R2: 0.723

[1] "Linear␣model␣A+B"
AIC index: 105.0284 BIC index: 108.5899 R2: 0.7493

*** Difference between models is weak

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
slope      2 438.84  219.42 11.9171 0.0004427 ***
distance   2  87.88   43.94  2.3866 0.1189442
surface    2 162.64   81.32  4.4166 0.0265959 *
Residuals 19 349.83   18.41
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.6634 Sigma: 1.991

[0] Testing each couple of factors
Multiplying model AxB
Linear model A+C
Linear model B+C

[1] "Mixed␣model␣C+AxB"
AIC index: 150.0427 BIC index: 158.5219 R2: 0.6999
```

Listing E.10: Experiment 1: Results for the 10th subject

```
# ---------------
# Subject:  10
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df  Sum Sq Mean Sq  F value   Pr(>F)
slope               2  631.44  315.72 437.1538 0.002282 **
distance            2   45.44   22.72   1.5792 0.387725
rep                 1    5.56    5.56
slope:distance      4   92.22   23.06   3.0515 0.152692
slope:rep           2    1.44    0.72
distance:rep        2   28.78   14.39
slope:distance:rep  4   30.22    7.56
Residuals           0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.921 F(8,9): 13.11 AIC: 94.469 BIC: 103.373

[1] "Multiplying␣model␣AxB"
AIC index: 96.8967 BIC index: 99.56781 R2: 0.8017

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
slope      2 1527.27  763.63  56.751 9.706e-09 ***
distance   2  627.00  313.50  23.299 7.723e-06 ***
surface    2  619.57  309.78  23.022 8.369e-06 ***
Residuals 19  255.66   13.46
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.9156 Sigma: 3.091

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C

[1] "Mixed␣model␣AxC+B"
AIC index: 148.7955 BIC index: 153.2747 R2: 0.9169

[1] "Linear␣model␣A+B+C"
AIC index: 147.7991 BIC index: 154.2783 R2: 0.9137

*** Difference among models is weak
```

Listing E.11: Experiment 1: Results for the 11th subject

```
# ---------------
# Subject:  11
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df Sum Sq Mean Sq F value  Pr(>F)
slope               2 668.11  334.06 11.7671 0.07833 .
distance            2  65.44   32.72 12.0204 0.07680 .
rep                 1  20.06   20.06
slope:distance      4  12.89    3.22  0.3766 0.81641
slope:rep           2  56.78   28.39
distance:rep        2   5.44    2.72
slope:distance:rep  4  34.22    8.56
Residuals           0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.865 F(8,9): 7.208 AIC: 104.697 BIC: 113.601

[1] "Linear␣model␣A+B"
AIC index: 95.70295 BIC index: 99.26444 R2: 0.8339

[1] "Multiplying␣model␣AxB"
AIC index: 97.79057 BIC index: 100.4617 R2: 0.7771

*** Difference between models is weak

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value   Pr(>F)
slope      2 374.88  187.44  8.5779 0.002215 **
distance   2 119.14   59.57  2.7261 0.091018 .
surface    2 336.61  168.30  7.7021 0.003551 **
Residuals 19 415.18   21.85
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.6667 Sigma: 3.023

[0] Testing each couple of factors
Linear model A+B (weak)
Linear model A+C
Linear model B+C

[1] "Linear␣model␣A+B+C"
AIC index: 161.5919 BIC index: 168.0711 R2: 0.6482

[1] "Mixed␣model␣AxC+B"
AIC index: 164.0818 BIC index: 168.5610 R2: 0.6441

*** Difference among models is weak
```

Listing E.12: Experiment 1: Results for the 12th subject

```
# ---------------
# Subject:  12
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df  Sum Sq Mean Sq F value  Pr(>F)
slope               2 135.111  67.556 10.8571 0.08434 .
distance            2 252.111 126.056 37.1967 0.02618 *
rep                 1  14.222  14.222
slope:distance      4  68.556  17.139  3.3351 0.13508
slope:rep           2  12.444   6.222
distance:rep        2   6.778   3.389
slope:distance:rep  4  20.556   5.139
Residuals           0   0.000
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.894 F(8,9): 9.495 AIC: 90.857 BIC: 99.761

[1] "Linear␣model␣A+B"
AIC index: 94.14808 BIC index: 97.70956 R2: 0.7517
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
slope      2 392.36  196.18  25.284 4.419e-06 ***
distance   2 183.05   91.52  11.796 0.0004673 ***
surface    2 416.33  208.17  26.829 2.924e-06 ***
Residuals 19 147.42    7.76
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.8706 Sigma: 2.982

[0] Testing each couple of factors
Linear model A+B (weak)
Linear model A+C
Linear model B+C

[1] "Linear␣model␣A+B+C"
AIC index: 142.4044 BIC index: 148.8836 R2: 0.8141

[1] "Mixed␣model␣AxC+B"
AIC index: 144.4916 BIC index: 148.9708 R2: 0.8134

*** Difference among models is weak
```

Listing E.13: Experiment 1: Results for the 13th subject

```
# ---------------
# Subject:  13
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df Sum Sq Mean Sq  F value   Pr(>F)
slope               2 432.11  216.06  11.8930 0.077562 .
distance            2 143.44   71.72 430.3333 0.002318 **
rep                 1   0.50    0.50
slope:distance      4  41.22   10.31   1.2006 0.431807
slope:rep           2  36.33   18.17
distance:rep        2   0.33    0.17
slope:distance:rep  4  34.33    8.58
Residuals           0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.896 F(8,9): 9.705 AIC: 95.91 BIC: 104.813

[1] "Multiplying␣model␣AxB"
AIC index: 87.22562 BIC index: 89.89673 R2: 0.8333
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
slope      2 1285.32  642.66 35.1565 4.116e-07 ***
distance   2  262.69  131.35  7.1853  0.004745 **
surface    2  629.10  314.55 17.2074 5.437e-05 ***
Residuals 19  347.32   18.28
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.8624 Sigma: 1.92

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C

[1] "Mixed␣model␣AxC+B"
AIC index: 154.3803 BIC index: 158.8595 R2: 0.8984
```

Listing E.14: Experiment 1: Results for the 14th subject

```
# ---------------
# Subject:  14
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                  Df Sum Sq Mean Sq F value  Pr(>F)
slope              2 534.78  267.39 76.3968 0.01292 *
distance           2  70.78   35.39  4.9380 0.16841
rep                1   0.50    0.50
slope:distance     4  52.89   13.22  1.9833 0.26175
slope:rep          2   7.00    3.50
distance:rep       2  14.33    7.17
slope:distance:rep 4  26.67    6.67
Residuals          0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.931 F(8,9): 15.273 AIC: 88.923 BIC: 97.827

[1] "Multiplying␣model␣AxB"
AIC index: 88.40501 BIC index: 91.07613 R2: 0.803
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
slope      2 688.15  344.07 32.3201 7.677e-07 ***
distance   2 196.06   98.03  9.2084  0.001600 **
surface    2 353.67  176.83 16.6105 6.739e-05 ***
Residuals 19 202.27   10.65
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.8595 Sigma: 2.085

[0] Testing each couple of factors
Multiplying model AxB (weak)
Linear model A+C
Linear model B+C

[1] "Mixed␣model␣C+AxB"
AIC index: 139.1617 BIC index: 147.6409 R2: 0.86

[1] "Mixed␣model␣AxC+B"
AIC index: 143.9467 BIC index: 148.4259 R2: 0.8581

*** Difference among models is weak
```

Listing E.15: Experiment 1: Results for the 15th subject

```
# ---------------
# Subject:  15
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                  Df  Sum Sq  Mean Sq  F value   Pr(>F)
slope              2  403.00   201.50  45.9114  0.02132 *
distance           2    4.00     2.00   2.2500  0.30769
rep                1    0.22     0.22
slope:distance     4    3.00     0.75   0.1189  0.96850
slope:rep          2    8.78     4.39
distance:rep       2    1.78     0.89
slope:distance:rep 4   25.22     6.31
Residuals          0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.919 F(8,9): 12.812 AIC: 83.558 BIC: 92.462

[1] "Linear␣model␣A+B"
AIC index: 76.11138 BIC index: 79.67286 R2: 0.8949
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq  Mean Sq  F value     Pr(>F)
slope      2  536.65   268.32  53.5125  1.563e-08 ***
distance   2   24.91    12.46   2.4841   0.110058
surface    2   75.62    37.81   7.5408   0.003884 **
Residuals 19   95.27     5.01
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.8699 Sigma: 0.798

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C (weak)

[1] "Mixed␣model␣AxC+B"
AIC index: 129.0625 BIC index: 133.5417 R2: 0.8385

[1] "Linear␣model␣A+B+C"
AIC index: 127.9438 BIC index: 134.4230 R2: 0.8359

*** Difference among models is weak
```

Listing E.16: Experiment 1: Results for the 16th subject

```
# ---------------
# Subject:  16
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df  Sum Sq Mean Sq F value  Pr(>F)
slope               2 1167.44  583.72 18.2097 0.05206 .
distance            2  155.44   77.72  5.6640 0.15006
rep                 1   22.22   22.22
slope:distance      4   43.56   10.89  0.1828 0.93571
slope:rep           2   64.11   32.06
distance:rep        2   27.44   13.72
slope:distance:rep  4  238.22   59.56
Residuals           0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.795 F(8,9): 4.367 AIC: 124.6 BIC: 133.504

[1] "Multiplying␣model␣AxB"
AIC index: 113.7276 BIC index: 116.3987 R2: 0.756

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
slope      2 1509.80  754.90 31.3583 9.577e-07 ***
distance   2  346.71  173.36  7.2011  0.004702 **
surface    2  408.51  204.25  8.4846  0.002327 **
Residuals 19  457.39   24.07
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.832 Sigma: 3.711

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C

[1] "Mixed␣model␣AxC+B"
AIC index: 163.7544 BIC index: 168.2336 R2: 0.8445

[1] "Linear␣model␣A+B+C"
AIC index: 162.2915 BIC index: 168.7706 R2: 0.8347

*** Difference among models is weak
```

Listing E.17: Experiment 1: Results for the 17th subject

```
# ---------------
# Subject:  17
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df Sum Sq Mean Sq F value  Pr(>F)
slope               2 937.33  468.67 78.1111 0.01264 *
distance            2  70.33   35.17  4.3061 0.18846
rep                 1   0.50    0.50
slope:distance      4  16.33    4.08  0.2344 0.90549
slope:rep           2  12.00    6.00
distance:rep        2  16.33    8.17
slope:distance:rep  4  69.67   17.42
Residuals           0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.912 F(8,9): 11.695 AIC: 101.676 BIC: 110.58

[1] "Linear␣model␣A+B"
AIC index: 92.79183 BIC index: 96.35331 R2: 0.8954

[1] "Multiplying␣model␣AxB"
AIC index: 94.1257 BIC index: 96.79682 R2: 0.8759

*** Difference between models is weak

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq  F value    Pr(>F)
slope      2 1664.29  832.14 205.2101 1.367e-13 ***
distance   2   79.31   39.65   9.7787  0.001203 **
surface    2  270.39  135.20  33.3401 6.106e-07 ***
Residuals 19   77.05    4.06
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.9632 Sigma: 2.059

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C (weak)

[1] "Mixed␣model␣AxC+B"
AIC index: 121.3279 BIC index: 125.8070 R2: 0.9566

[1] "Linear␣model␣A+B+C"
AIC index: 119.4723 BIC index: 125.9514 R2: 0.9558

*** Difference among models is weak
```

Listing E.18: Experiment 1: Results for the 18th subject

```
# ---------------
# Subject:  18
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df  Sum Sq Mean Sq F value  Pr(>F)
slope               2 1303.00  651.50  5.0744 0.16462
distance            2  586.33  293.17 32.3742 0.02996 *
rep                 1  430.22  430.22
slope:distance      4  120.67   30.17  3.6689 0.11797
slope:rep           2  256.78  128.39
distance:rep        2   18.11    9.06
slope:distance:rep  4   32.89    8.22
Residuals           0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.731 F(8,9): 3.064 AIC: 137.926 BIC: 146.83

[1] "Multiplying␣model␣AxB"
AIC index: 128.5163 BIC index: 131.1874 R2: 0.6831

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
slope      2 2838.60 1419.30 87.2919 2.648e-10 ***
distance   2  157.63   78.81  4.8473 0.0199092 *
surface    2  441.25  220.63 13.5692 0.0002186 ***
Residuals 19  308.93   16.26
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.9175 Sigma: 4.052

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C (weak)

[1] "Mixed␣model␣AxC+B"
AIC index: 158.4493 BIC index: 162.9285 R2: 0.9052

[1] "Linear␣model␣A+B+C"
AIC index: 156.4495 BIC index: 162.9287 R2: 0.9052

*** Difference among models is weak
```

Listing E.19: Experiment 1: Results for the 19th subject

```
# ---------------
# Subject:  19
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df Sum Sq Mean Sq F value  Pr(>F)
slope               2 382.11  191.06 17.2814 0.05470 .
distance            2 494.78  247.39 11.9383 0.07729 .
rep                 1  22.22   22.22
slope:distance      4 103.22   25.81  4.0925 0.10054
slope:rep           2  22.11   11.06
distance:rep        2  41.44   20.72
slope:distance:rep  4  25.22    6.31
Residuals           0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.898 F(8,9): 9.934 AIC: 103.827 BIC: 112.73

[1] "Linear␣model␣A+B"
AIC index: 104.4259 BIC index: 107.9874 R2: 0.7823
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
slope      2 1538.02  769.01  34.697 4.541e-07 ***
distance   2 1072.08  536.04  24.185 5.993e-06 ***
surface    2 1743.77  871.88  39.338 1.759e-07 ***
Residuals 19  421.11   22.16
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.9118 Sigma: 4.101

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C

[1] "Mixed␣model␣AxC+B"
AIC index: 190.3989 BIC index: 194.8781 R2: 0.7536

[1] "Linear␣model␣A+B+C"
AIC index: 188.5081 BIC index: 194.9873 R2: 0.7526

*** Difference among models is weak
```

Listing E.20: Experiment 1: Results for the 20th subject

```
# ---------------
# Subject:  20
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df  Sum Sq Mean Sq F value  Pr(>F)
slope               2 2283.44 1141.72 28.5035 0.03389 *
distance            2  990.11  495.06 17.8577 0.05303 .
rep                 1   53.39   53.39
slope:distance      4  272.89   68.22  2.0131 0.25733
slope:rep           2   80.11   40.06
distance:rep        2   55.44   27.72
slope:distance:rep  4  135.56   33.89
Residuals           0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.916 F(8,9): 12.295 AIC: 123.136 BIC: 132.04

[1] "Multiplying␣model␣AxB"
AIC index: 119.4774 BIC index: 122.1486 R2: 0.8411

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
slope      2 1937.14  968.57  20.957 1.561e-05 ***
distance   2  270.83  135.42   2.930   0.07779 .
surface    2 2698.30 1349.15  29.192 1.607e-06 ***
Residuals 19  878.13   46.22
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.8482 Sigma: 5.499

[0] Testing each couple of factors
Multiplying model AxB (weak)
Linear model A+C
Linear model B+C (weak)

[1] "Mixed␣model␣AxC+B"
AIC index: 184.8931 BIC index: 189.3723 R2: 0.8319

[1] "Mixed␣model␣C+AxB"
AIC index: 181.4931 BIC index: 189.9723 R2: 0.8254

*** Difference among models is weak
```

Listing E.21: Experiment 1: Results for the 21th subject

```
# ---------------
# Subject:  21
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df  Sum Sq  Mean Sq  F value   Pr(>F)
slope               2  836.11   418.06  80.9140  0.01221 *
distance            2  120.78    60.39  13.4198  0.06935 .
rep                 1    0.50     0.50
slope:distance      4   70.89    17.72   6.6458  0.04684 *
slope:rep           2   10.33     5.17
distance:rep        2    9.00     4.50
slope:distance:rep  4   10.67     2.67
Residuals           0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.971 F(8,9): 37.91 AIC: 80.574 BIC: 89.478

[1] "Multiplying␣model␣AxB"
AIC index: 83.4 BIC index: 86.07111 R2: 0.9262
*** Positive evidences for the model

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value     Pr(>F)
slope      2  796.41  398.20 90.2596  1.987e-10 ***
distance   2   79.70   39.85  9.0328   0.001750 **
surface    2  450.11  225.06 51.0127  2.295e-08 ***
Residuals 19   83.82    4.41
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.9406 Sigma: 1.152

[0] Testing each couple of factors
Linear model A+B (weak)
Linear model A+C
Linear model B+C (weak)

[1] "Linear␣model␣A+B+C"
AIC index: 126.8536 BIC index: 133.3328 R2: 0.918

[1] "Mixed␣model␣AxC+B"
AIC index: 129.0071 BIC index: 133.4863 R2: 0.9187

*** Difference among models is weak
```

Listing E.22: Experiment 1: Results for the 22th subject

```
# ---------------
# Subject:  22
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df Sum Sq Mean Sq F value Pr(>F)
slope               2 521.44  260.72  7.9677 0.1115
distance            2 314.78  157.39  5.0680 0.1648
rep                 1  34.72   34.72
slope:distance      4  54.89   13.72  0.2426 0.9005
slope:rep           2  65.44   32.72
distance:rep        2  62.11   31.06
slope:distance:rep  4 226.22   56.56
Residuals           0   0.00

Fit Indexes for model without repeted factor
 R2: 0.696 F(8,9): 2.58 AIC: 126.376 BIC: 135.28

[1] "Multiplying␣model␣AxB"
AIC index: 118.3831 BIC index: 121.0542 R2: 0.5324

[1] "Linear␣model␣A+B"
AIC index: 118.8774 BIC index: 122.4389 R2: 0.5492

*** Difference between models is weak

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
slope      2 241.11  120.56  9.4412 0.0014224 **
distance   2 333.82  166.91 13.0710 0.0002689 ***
surface    2 243.83  121.92  9.5476 0.0013486 **
Residuals 19 242.62   12.77
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.7714 Sigma: 2.164

[0] Testing each couple of factors
Linear model A+B
Linear model A+C (weak)
Linear model B+C

[1] "Mixed␣model␣AxC+B"
AIC index: 150.6460 BIC index: 155.1252 R2: 0.7451
```

Listing E.23: Experiment 1: Results for the 23th subject

```
# ---------------
# Subject:  23
# ---------------
Analisi Trial
Analysis of Variance Table

Response: y
                   Df Sum Sq Mean Sq F value Pr(>F)
slope               2 424.78  212.39  1.5598 0.3907
distance            2 112.11   56.06  1.2986 0.4351
rep                 1 722.00  722.00
slope:distance      4 107.56   26.89  0.5238 0.7267
slope:rep           2 272.33  136.17
distance:rep        2  86.33   43.17
slope:distance:rep  4 205.33   51.33
Residuals           0   0.00

Fit Indexes for model without repeted factor
 R2: 0.334 F(8,9): 0.564 AIC: 147.922 BIC: 156.826

[1] "Multiplying␣model␣AxB"
AIC index: 135.3291 BIC index: 138.0002 R2: 0.2819

Disegno sperimentale
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
slope      2  483.18  241.59  12.478 0.0003464 ***
distance   2  626.98  313.49  16.191 7.860e-05 ***
surface    2 1656.70  828.35  42.783 9.204e-08 ***
Residuals 19  367.87   19.36
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.8826 Sigma: 4.878

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C

[1] "Linear␣model␣A+B+C"
AIC index: 158.0781 BIC index: 164.5573 R2: 0.8803

[1] "Mixed␣model␣AxC+B"
AIC index: 160.1817 BIC index: 164.6609 R2: 0.8779

*** Difference among models is weak
```

Listing E.24: Experiment 1: Results for the Trial step

```
Analysis of Variance Table

Response: y
                      Df  Sum Sq Mean Sq  F value    Pr(>F)
distance               2  2836.0  1418.0  90.5579 0.0109220 *
slope                  2 11359.4  5679.7 204.6125 0.0048635 **
id                    19 18425.5   969.8  20.9744 5.918e-09 ***
rep                    1   0.025   0.025
distance:slope         4   135.4    33.9   5.3382 0.0668236 .
distance:id           38  1154.2    30.4   4.4735 5.471e-06 ***
slope:id              38  1660.8    43.7   3.0061 0.0004992 ***
distance:rep           2    31.3    15.7
slope:rep              2    55.5    27.8
id:rep                19   878.5    46.2
distance:slope:id     76   783.0    10.3   1.0766 0.3741994
distance:slope:rep     4    25.4     6.3
distance:id:rep       38   258.0     6.8
slope:id:rep          38   552.5    14.5
distance:slope:id:rep 76   727.3     9.6
Residuals              0     0.0
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.935 F(179,180): 14.458 AIC: 2085.376 BIC: 2788.76

[1] "Multiplying␣model␣AxB"
AIC index: 2553.087 BIC index: 2564.745 R2: 0.347
*** Positive evidences for the model
```

Listing E.25: Experiment 1: Results for the Experimental step

```
Analysis of Variance Table

Response: y
                        Df  Sum Sq Mean Sq  F value  Pr(>F)
distance                 2  3697.3  1848.6  66.0305 < 2e-16 ***
slope                    2 15753.7  7876.8 281.3471 < 2e-16 ***
surface                  2  8641.1  4320.6 154.3231 < 2e-16 ***
id                      19 27331.5  1438.5  51.3808 < 2e-16 ***
distance:slope           4    66.1    16.5   0.5902 0.66988
distance:surface         4   272.2    68.1   2.4310 0.04677 *
slope:surface            4    63.8    15.9   0.5695 0.68486
distance:slope:surface   8    34.9     4.4   0.1557 0.99611
Residuals              493 13802.4    28.0
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1
R-square: 0.8019 Sigma: 3.467

[0] Testing each couple of factors
Linear model A+B
Linear model A+C
Linear model B+C

[1] "Linear␣model␣A+B+C"
AIC index: 3909.647 BIC index: 3931.105 R2: 0.4084

[1] "Mixed␣model␣AxC+B"
AIC index: 3911.761 BIC index: 3931.219 R2: 0.4087

*** Difference among models is weak
```

Listing E.26: Experiment 2: Model selection and estimated values for the single-subject analysis (1 of 2).

```
####################################################
1  Predictability
####################################################

Analysis of Variance Table

Response: y
              Df Sum Sq Mean Sq F value    Pr(>F)
livelli        2 925.00  462.50  1067.3 2.204e-12 ***
trial          5   3.17    0.63
livelli:trial 10   4.33    0.43
Residuals      0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.992 F(2,15): 925 AIC: 43.323 BIC: 46.885

Optimal selection between models:  Equal - Bounded
Estimation mode: B

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    9.000    0.236   8.498   9.502  38.184 0.0000 ***
2           e    9.000    0.289   8.385   9.615  31.177 0.0000 ***

     SSreg  SSres  Rsquare        F dfk dfres       p  AIC      BIC
1 2430.000 15.000 0.993865 1215.000   2    15 0.00000 53.8 56.47111

####################################################
2  Honesty
####################################################

Analysis of Variance Table

Response: y
              Df  Sum Sq Mean Sq F value    Pr(>F)
livelli        2 1009.33  504.67  315.42 9.253e-10 ***
trial          5    0.67    0.13
livelli:trial 10   16.00    1.60
Residuals      0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.984 F(2,15): 454.2 AIC: 57.696 BIC: 61.258

Optimal selection between models:  Equal - Bounded
Estimation mode: B

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)   10.500    0.269   9.927  11.073  39.071 0.0000 ***
2           e    9.500    0.329   8.798  10.202  28.863 0.0000 ***

     SSreg  SSres   Rsquare        F dfk dfres       p      AIC      BIC
1 3067.500 19.500 0.9936832 1179.808   2    15 0.00000 58.52256 61.19367
```

Listing E.27: Experiment 2: Model selection and estimated values for the single-subject analysis (2 of 2).

```
####################################################
3  Benevolence
####################################################

Analysis of Variance Table

Response: y
              Df  Sum Sq Mean Sq F value    Pr(>F)
livelli        2 1045.33  522.67  186.67 1.208e-08 ***
trial          5   17.17    3.43
livelli:trial 10   28.00    2.80
Residuals      0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.959 F(2,15): 173.579 AIC: 75.642 BIC: 79.203

Optimal selection between models:  Equal - Bounded
Estimation mode: B

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)   10.000    0.451   9.038  10.962  22.156 0.0000 ***
2           e   10.000    0.553   8.822  11.178  18.091 0.0000 ***

     SSreg  SSres   Rsquare       F dfk dfres       p     AIC      BIC
1 3000.000 55.000 0.9819967 409.091   2    15 0.00000 77.1871 79.85821

####################################################
4  Competence
####################################################

Analysis of Variance Table

Response: y
              Df  Sum Sq Mean Sq F value    Pr(>F)
livelli        2 1064.33  532.17  409.36 2.558e-10 ***
trial          5    7.17    1.43
livelli:trial 10   13.00    1.30
Residuals      0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.981 F(2,15): 395.826 AIC: 61.128 BIC: 64.689

Optimal selection between models:  Equal - Bounded
Estimation mode: B

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)   10.000    0.279   9.406  10.594  35.857 0.0000 ***
2           e    9.500    0.342   8.772  10.228  27.813 0.0000 ***

     SSreg  SSres   Rsquare        F dfk dfres       p     AIC      BIC
1 2883.000 21.000 0.9927686 1029.643   2    15 0.00000 59.8565 62.52761
```

Listing E.28: Experiment 2: Model selection and estimated weights for the single-subject analysis (1 of 3).

```
####################################################
1  Predictability Honesty
####################################################

Analysis of Variance Table

Response: y
            Df Sum Sq Mean Sq F value    Pr(>F)
F1           2 525.35  262.68 113.753 1.323e-07 ***
F2           2 920.61  460.30 175.430 1.634e-08 ***
trial        5  33.05    6.61
F1:F2        4  58.31   14.58  11.022 6.904e-05 ***
F1:trial    10  23.09    2.31
F2:trial    10  26.24    2.62
F1:F2:trial 20  26.45    1.32
Residuals    0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.933 F(8,45): 77.747 AIC: 211.09 BIC: 230.98
        AdjR2   RSS df   AIC   BIC    c0                               Model AIC BIC
RLS      0.97 174.7  3 224.6 232.6 -1.17 0.42 0.57  NaN  NaN  NaN  NaN a+b=1
w A=B    0.93 287.0  3 251.5 259.4  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.95 188.1  3 228.6 236.6 -0.68 1.00 1.00 1.00 1.00 1.00 1.00
w x6     0.96 115.9  5 206.5 218.4 -0.31 1.00 1.00 1.00 0.71 1.55 2.09       AIC BIC
AVERAGE  0.95 166.6  5 226.1 238.0 -0.46 1.00 1.75 1.75 2.09 1.75 1.75
conj.    0.92 342.0  5 264.9 276.9 -0.94 2.00 2.00 1.00 3.00 3.00 2.00
AxB      0.93 357.3  2 261.3 267.3  4.39 0.84  NaN  NaN  NaN  NaN  NaN

####################################################
2  Predictability Benevolence
####################################################

Analysis of Variance Table

Response: y
            Df Sum Sq Mean Sq F value    Pr(>F)
F1           2 572.76  286.38 259.645 2.407e-09 ***
F2           2 674.40  337.20 131.453 6.606e-08 ***
trial        5  48.67    9.73
F1:F2        4  83.23   20.81  12.759 2.562e-05 ***
F1:trial    10  11.03    1.10
F2:trial    10  25.65    2.57
F1:F2:trial 20  32.61    1.63
Residuals    0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.919 F(8,45): 63.437 AIC: 215.442 BIC: 235.332

        AdjR2   RSS df   AIC   BIC    c0                               Model AIC BIC
RLS      0.96 208.8  3 234.3 242.2 -0.42 0.50 0.40  NaN  NaN  NaN  NaN a+b=1
w A=B    0.92 354.5  3 262.9 270.8  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.94 207.2  3 233.9 241.8 -0.83 1.00 1.00 1.00 1.00 1.00 1.00
w x6     0.95 127.5  6 213.6 227.6 -0.37 1.00 3.00 2.50 1.50 2.00 2.00       AIC BIC
AVERAGE  0.94 163.4  4 223.0 233.0 -0.60 1.00 1.39 1.39 1.39 1.39 1.39
conj.    0.90 341.0  4 262.8 272.7 -0.84 2.00 2.00 1.00 2.00 2.00 2.00
AxB      0.94 292.0  2 250.4 256.3  4.35 0.79  NaN  NaN  NaN  NaN  NaN
```

Listing E.29: Experiment 2: Model selection and estimated weights for the single-subject analysis (2 of 3).

```
####################################################
3  Predictability Competence
####################################################

Analysis of Variance Table

Response: y
            Df Sum Sq Mean Sq F value    Pr(>F)
F1           2 413.82  206.91 163.256 2.317e-08 ***
F2           2 820.18  410.09 106.650 1.801e-07 ***
trial        5  11.78    2.36
F1:F2        4  42.80   10.70   6.163  0.002119 **
F1:trial    10  12.67    1.27
F2:trial    10  38.45    3.85
F1:F2:trial 20  34.73    1.74
Residuals    0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.929 F(8,45): 73.561 AIC: 205.226 BIC: 225.116

        AdjR2   RSS df   AIC   BIC    c0                                Model AIC  BIC
RLS      0.97 136.6  3 211.4 219.3 -0.57 0.42 0.49  NaN  NaN  NaN  NaN a+b=1      BIC*
w A=B    0.92 304.7  3 254.7 262.6  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.95 163.9  3 221.2 229.1 -0.81 1.00 1.00 1.00 1.00 1.00 1.00
w x6     0.96 111.2  6 206.2 220.2 -0.58 1.00 2.00 1.50 1.50 2.00 2.00       AIC
AVERAGE  0.96 139.5  5 216.5 228.4 -0.73 1.00 1.85 1.85 1.93 1.85 1.85
conj.    0.93 247.3  5 247.4 259.4 -1.03 2.00 2.00 1.00 3.00 3.00 2.00
AxB      0.92 356.9  2 261.2 267.2  4.30 0.79  NaN  NaN  NaN  NaN  NaN

####################################################
4  Honesty Benevolence
####################################################

Analysis of Variance Table

Response: y
            Df  Sum Sq Mean Sq  F value    Pr(>F)
F1           2 1024.00  512.00 1440.000 4.960e-13 ***
F2           2  589.35  294.68  239.357 3.587e-09 ***
trial        5    9.78    1.96
F1:F2        4   54.70   13.68   12.768 2.550e-05 ***
F1:trial    10    3.56    0.36
F2:trial    10   12.31    1.23
F1:F2:trial 20   21.42    1.07
Residuals    0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.973 F(8,45): 199.351 AIC: 165.825 BIC: 185.715

        AdjR2   RSS df   AIC   BIC    c0                                Model AIC BIC
RLS      0.98  89.2  3 188.3 196.3 -1.37 0.58 0.42  NaN  NaN  NaN  NaN a+b=1
w A=B    0.95 231.7  3 239.9 247.9  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.98  92.1  3 190.1 198.0 -0.71 1.50 1.50 1.50 1.00 1.00 1.00
w x6     0.98  59.1  5 170.2 182.1 -0.43 1.50 2.00 2.00 1.00 1.50 1.50       AIC BIC
AVERAGE  0.97  98.6  4 195.8 205.7 -0.31 1.00 1.11 1.11 1.11 1.11 1.11
conj.    0.97 124.4  4 208.3 218.3 -0.66 1.50 1.50 1.00 1.00 1.00 1.00
AxB      0.94 339.2  2 258.5 264.4  4.81 0.79  NaN  NaN  NaN  NaN  NaN
```

Listing E.30: Experiment 2: Model selection and estimated weights for the single-subject analysis (3 of 3).

```
####################################################
5  Honesty Competence
####################################################

Analysis of Variance Table

Response: y
            Df Sum Sq Mean Sq F value    Pr(>F)
F1           2 810.04  405.02 133.197 6.199e-08 ***
F2           2 697.93  348.96 128.365 7.407e-08 ***
trial        5  35.70    7.14
F1:F2        4  80.19   20.05  12.014 3.872e-05 ***
F1:trial    10  30.41    3.04
F2:trial    10  27.19    2.72
F1:F2:trial 20  33.37    1.67
Residuals    0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.926 F(8,45): 70.526 AIC: 219.284 BIC: 239.174

        AdjR2   RSS df   AIC   BIC    c0                                    Model AIC BIC
RLS      0.94 422.6  3 272.3 280.3 -0.71 0.69 0.35  NaN  NaN  NaN  NaN a+b=1
w A=B    0.93 291.9  3 252.4 260.3  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.94 212.4  3 235.2 243.2 -0.61 1.00 1.00 1.00 1.00 1.00 1.00
w x6     0.96 132.5  6 215.7 229.6 -0.08 1.00 2.00 2.00 1.00 1.50 2.00       AIC BIC
AVERAGE  0.94 202.5  3 232.6 240.6 -0.09 1.00 1.00 1.00 1.00 1.00 1.00
conj.    0.92 324.5  5 262.1 274.0 -0.75 1.50 1.50 1.00 1.50 1.50 1.00
AxB      0.95 328.9  2 256.8 262.8  5.12 0.79  NaN  NaN  NaN  NaN  NaN

####################################################
6  Benevolence Competence
####################################################

Analysis of Variance Table

Response: y
            Df Sum Sq Mean Sq F value    Pr(>F)
F1           2 541.03  270.52  81.047 6.625e-07 ***
F2           2 952.16  476.08 206.593 7.368e-09 ***
trial        5  26.08    5.22
F1:F2        4  53.48   13.37  14.748 9.170e-06 ***
F1:trial    10  33.38    3.34
F2:trial    10  23.04    2.30
F1:F2:trial 20  18.13    0.91
Residuals    0   0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.939 F(8,45): 86.453 AIC: 206.86 BIC: 226.75

        AdjR2   RSS df   AIC   BIC    c0                                    Model AIC BIC
RLS      0.97 163.2  3 221.0 228.9 -0.44 0.44 0.53  NaN  NaN  NaN  NaN a+b=1
w A=B    0.94 250.6  3 244.1 252.1  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.96 160.2  3 220.0 227.9 -0.48 1.00 1.00 1.00 1.50 1.50 1.50
w x6     0.97 106.2  6 203.7 217.7 -0.13 1.00 2.50 2.00 2.00 2.00 2.50       AIC BIC
AVERAGE  0.96 144.1  5 218.2 230.2 -0.11 1.00 1.11 1.11 1.19 1.11 1.11
conj.    0.95 198.7  4 233.6 243.6 -0.43 1.00 1.00 1.00 1.50 1.50 1.00
AxB      0.93 392.3  2 266.3 272.3  5.30 0.76  NaN  NaN  NaN  NaN  NaN
```

Listing E.31: Experiment 2: Estimated values for the factor *predictability*. All the sample.

```
#####################################################
1  Predictability
#####################################################

Analysis of Variance Table

Response: y
                  Df  Sum Sq Mean Sq F value    Pr(>F)
livelli            2 1741.58  870.79 99.6139 2.494e-07 ***
id                 3   10.50    3.50  0.3391  0.797390
trial              5    7.50    1.50
livelli:id         6  132.08   22.01  4.4150  0.002594 **
livelli:trial     10   87.42    8.74
id:trial          15  154.83   10.32
livelli:id:trial  30  149.58    4.99
Residuals          0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.825 F(11,60): 25.736 AIC: 353.673 BIC: 383.269

Optimal selection between models:  [3] Equal - Bounded

[1] Equal values model X=[-1, 0,1]

Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    8.889    0.265   8.361   9.416  33.605 0.0000 ***
2    object@x    6.612    0.374   5.866   7.359  17.665 0.0000 ***

     SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 7787.587 561.658 0.9327295 478.355   2    69 0.00000 358.2315 365.0615


[2] Different values model X=[1,0,0, 0,0,-1]

Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    8.497    0.494   7.512   9.482  17.207 0.0000 ***
2  object@xd1    7.216    0.822   5.576   8.857   8.775 0.0000 ***
3  object@xd2    6.042    0.612   4.822   7.262   9.880 0.0000 ***

     SSreg   SSres  Rsquare       F dfk dfres       p      AIC      BIC
1 7803.456 562.268 0.932789 314.580   3    68 0.00000 360.3097 369.4163


>> Bic difference between models:  Positive

[3] Approximate equal values model

Estimation mode: B

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    9.000    0.335   8.331   9.669  26.830 0.0000 ***
2           e    6.500    0.411   5.680   7.320  15.822 0.0000 ***

     SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 7860.000 559.000 0.9336026 485.098   2    69 0.00000 357.8900 364.7199


>> Bic difference between models:  Weak
```

Listing E.32: Experiment 2: Estimated values for the factor *honesty*. All the sample.

```
#####################################################
2  Honesty
#####################################################

Analysis of Variance Table

Response: y
                  Df  Sum Sq Mean Sq  F value     Pr(>F)
livelli            2  3374.2  1687.1 159.7042 2.578e-08 ***
id                 3    19.4     6.5   4.7684   0.01579 *
trial              5     2.6     0.5
livelli:id         6    43.1     7.2   3.1247   0.01686 *
livelli:trial     10   105.6    10.6
id:trial          15    20.4     1.4
livelli:id:trial  30    69.0     2.3
Residuals          0     0.0
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.946 F(11,60): 94.837 AIC: 303.041 BIC: 332.638

Optimal selection between models:  [4] Different - Bounded

[1] Equal values model X=[-1, 0,1]

Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    9.962    0.201   9.562  10.362  49.622 0.0000 ***
2    object@x    8.959    0.173   8.614   9.304  51.800 0.0000 ***

      SSreg   SSres   Rsquare        F dfk dfres       p      AIC      BIC
1 10998.171 304.439 0.9730647 1246.347   2    69 0.00000 314.1371 320.9671


[2] Different values model X=[1,0,0, 0,0,-1]

Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    9.113    0.324   8.467   9.759  28.141 0.0000 ***
2  object@xd1   10.203    0.456   9.293  11.113  22.374 0.0000 ***
3  object@xd2    7.643    0.438   6.769   8.518  17.434 0.0000 ***

      SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 10999.920 278.396 0.9753158 895.599   3    68 0.00000 309.6985 318.8052


>> Bic difference between models:  Positive

[4] Approximate different values model

Estimation mode: B

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    9.000    0.405   8.191   9.809  22.209 0.0000 ***
2          d1   10.000    0.573   8.856  11.144  17.449 0.0000 ***
3          d2    7.500    0.573   6.356   8.644  13.087 0.0000 ***

      SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 10662.000 268.000 0.9754803 901.761   3    68 0.00000 306.9583 316.0649


>> Bic difference between models:  Positive
```

Listing E.33: Experiment 2: Estimated values for the factor *benevolence*. All the sample.

```
######################################################
3  Benevolence
######################################################

Analysis of Variance Table

Response: y
                  Df  Sum Sq Mean Sq  F value    Pr(>F)
livelli            2 3072.11 1536.06 143.4821  4.33e-08 ***
id                 3   98.04   32.68   7.4227 0.0028245 **
trial              5   18.57    3.71
livelli:id         6   88.67   14.78   5.8720 0.0003841 ***
livelli:trial     10  107.06   10.71
id:trial          15   66.04    4.40
livelli:id:trial  30   75.50    2.52
Residuals          0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.924 F(11,60): 66.533 AIC: 324.734 BIC: 354.331

Optimal selection between models:  [3] Equal - Bounded

[1] Equal values model X=[-1, 0,1]

Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    9.377    0.314   8.751  10.002  29.896 0.0000 ***
2    object@x    8.390    0.290   7.811   8.969  28.917 0.0000 ***

     SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 9709.391 462.201 0.9545596 724.737   2    69 0.00000 344.1992 351.0292


[2] Different values model X=[1,0,0, 0,0,-1]

Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    8.964    0.568   7.831  10.097  15.787 0.0000 ***
2  object@xd1    9.295    0.674   7.951  10.639  13.795 0.0000 ***
3  object@xd2    7.898    0.724   6.454   9.342  10.909 0.0000 ***

     SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 9957.436 478.997 0.9541034 471.197   3    68 0.00000 348.7692 357.8758


>> Bic difference between models:  Strong

[3] Approximate equal values model

Estimation mode: B

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    9.500    0.308   8.886  10.114  30.886 0.0000 ***
2           e    8.500    0.377   7.748   9.252  22.564 0.0000 ***

     SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 9966.000 470.000 0.9549636 731.547   2    69 0.00000 345.4039 352.2339


>> Bic difference between models:  Weak
```

Listing E.34: Experiment 2: Estimated values for the factor *competence*. All the sample.

```
#####################################################
4  Competence
#####################################################

Analysis of Variance Table

Response: y
                 Df  Sum Sq Mean Sq  F value    Pr(>F)
livelli           2 2545.44 1272.72 629.3681 3.042e-11 ***
id                3    3.15    1.05   0.3658    0.7787
trial             5    5.40    1.08
livelli:id        6  126.89   21.15   9.0491 1.125e-05 ***
livelli:trial    10   20.22    2.02
id:trial         15   43.10    2.87
livelli:id:trial 30   70.11    2.34
Residuals         0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.951 F(11,60): 105.116 AIC: 277.603 BIC: 307.2

Optimal selection between models:  [4] Different - Bounded

[1] Equal values model X=[-1, 0,1]

Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    9.658    0.275   9.109  10.207  35.092 0.0000 ***
2    object@x    7.249    0.368   6.514   7.984  19.672 0.0000 ***

     SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 9238.372 320.317 0.9664894 995.025   2    69 0.00000 317.7977 324.6277


[2] Different values model X=[1,0,0, 0,0,-1]

Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    8.475    0.263   7.951   9.000  32.214 0.0000 ***
2  object@xd1    9.346    0.584   8.181  10.510  16.008 0.0000 ***
3  object@xd2    5.380    0.540   4.302   6.457   9.961 0.0000 ***

     SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 9576.033 272.108 0.9723696 797.687   3    68 0.00000 308.0534 317.1601


>> Bic difference between models:  Strong

[4] Approximate different values model

Estimation mode: B

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    8.500    0.411   7.679   9.321  20.669 0.0000 ***
2          d1    9.500    0.582   8.339  10.661  16.335 0.0000 ***
3          d2    5.500    0.582   4.339   6.661   9.457 0.0000 ***

     SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 9726.000 276.000 0.9724055 798.754   3    68 0.00000 309.0761 318.1827


>> Bic difference between models:  Weak
```

Listing E.35: Experiment 2: Model selection and estimated weights for the couple of factors *predictability* × *honesty*. All the sample.

```
####################################################
1  Predictability Honesty
####################################################

Analysis of Variance Table

Response: y
               Df Sum Sq Mean Sq  F value    Pr(>F)
F1              2 1010.5   505.2  48.1912 7.339e-06 ***
F2              2 4210.0  2105.0 197.0522 9.280e-09 ***
id              3   98.0    32.7   6.2597 0.0057321 **
trial           5   60.4    12.1
F1:F2           4  192.1    48.0  14.8379 8.772e-06 ***
F1:id           6  136.1    22.7   4.2738 0.0031579 **
F2:id           6   61.6    10.3   3.3808 0.0114770 *
F1:trial       10  104.8    10.5
F2:trial       10  106.8    10.7
id:trial       15   78.3     5.2
F1:F2:id       12  129.1    10.8   3.9219 0.0001889 ***
F1:F2:trial    20   64.7     3.2
F1:id:trial    30  159.2     5.3
F2:id:trial    30   91.1     3.0
F1:F2:id:trial 60  164.6     2.7
Residuals       0    0.0
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.876 F(35,180): 36.172 AIC: 977.736 BIC: 1102.621
        AdjR2    RSS df    AIC    BIC    c0                                    Model AIC BIC
RLS      0.94 1534.8  3 1044.5 1058.0 -1.62 0.44 0.66  NaN  NaN  NaN  NaN a+b=1
w A=B    0.88 1813.0  3 1080.5 1094.0  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.90 1538.4  3 1045.0 1058.5 -0.33 1.00 1.00 1.00 1.50 1.50 1.50
w x6     0.90 1387.5  6 1028.7 1052.4 -0.00 1.00 1.50 1.00 1.11 2.50 3.00       AIC BIC
AVERAGE  0.91 1477.3  6 1042.3 1065.9 -0.31 1.00 1.96 1.96 2.43 2.12 1.96
conj.    0.88 1798.6  7 1086.8 1113.8 -0.45 1.55 2.30 1.00 1.55 2.30 1.00
AxB      0.90 2301.5  2 1130.1 1140.2  3.91 1.06  NaN  NaN  NaN  NaN  NaN

[regr] Additive model
Estimation mode: RLS

                     Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1            (Intercept)   -1.623    0.280  -2.175  -1.070  -5.787 0.0000 ***
2 object@xPredictability    0.442    0.034   0.375   0.509  13.006 0.0000 ***
3        object@xHonesty    0.663    0.024   0.615   0.711  27.284 0.0000 ***

      SSreg    SSres   Rsquare        F dfk dfres       p      AIC      BIC
1 23100.451 1534.835 0.9376977 1063.588   3   212 0.00000 1044.536 1058.037

[equal] Averaging model: overal equal weighted case
Estimation mode: W

             Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1             c0    0.000      NaN     NaN     NaN     NaN    NaN NaN
2 Predictability    0.500    0.027   0.448   0.552  18.851 0.0000 ***
3        Honesty    0.500    0.023   0.455   0.545  21.955 0.0000 ***

      SSreg    SSres  Rsquare       F dfk dfres       p      AIC      BIC
1 13891.500 1813.024 0.884554 541.452   3   212 0.00000 1080.516 1094.017

[Optim2] Averaging model: equal weighted case
Estimation mode: W

             Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1             c0   -0.334      NaN     NaN     NaN     NaN    NaN NaN
2 Predictability    1.000    0.024   0.952   1.048  40.927 0.0000 ***
3        Honesty    1.500    0.021   1.459   1.541  71.501 0.0000 ***

      SSreg    SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 13964.059 1538.449 0.9007613 641.421   3   212 0.00000 1045.044 1058.545

[Optim6] Averaging model: differential weighted case
Estimation mode: W

           Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1           c0   -0.002      NaN     NaN     NaN     NaN    NaN NaN
2 High Predic.    1.000    0.023   0.954   1.046  42.790 0.0000 ***
3 med. Predic.    1.500    0.020   1.460   1.540  74.755 0.0000 ***
4 Low  Predic.    1.000    0.023   0.954   1.046  42.790 0.0000 ***
5 High Honest.    1.109    0.020   1.069   1.148  55.248 0.0000 ***
6 med. Honest.    2.500    0.023   2.454   2.546 106.975 0.0000 ***
7 Low  Honest.    3.000    0.020   2.960   3.040 149.510 0.0000 ***

      SSreg    SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 12490.837 1387.524 0.9000225 313.578   6   209 0.00000 1028.741 1052.368

[Multiplying] Multiplying model
Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    3.906    0.364   3.188   4.625  10.718 0.0000 ***
2         AxB    1.059    0.040   0.980   1.139  26.242 0.0000 ***

      SSreg    SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 20999.241 2301.550 0.9012244 971.701   2   213 0.00000 1130.050 1140.176
```

Listing E.36: Experiment 2: Model selection and estimated weights for the couple of factors *predictability* × *benevolence*. All the sample.

```
##################################################
2  Predictability Benevolence
##################################################

Analysis of Variance Table

Response: y
               Df  Sum Sq Mean Sq  F value    Pr(>F)
F1              2   744.9   372.5 206.7137 7.347e-09 ***
F2              2  4079.6  2039.8 393.4652 3.111e-10 ***
id              3    67.4    22.5   4.2949 0.0224315 *
trial           5     9.7     1.9
F1:F2           4   253.9    63.5  35.3042 8.426e-09 ***
F1:id           6   155.0    25.8  14.7106 9.160e-08 ***
F2:id           6   122.3    20.4   6.7811 0.0001290 ***
F1:trial       10    18.0     1.8
F2:trial       10    51.8     5.2
id:trial       15    78.5     5.2
F1:F2:id       12    81.4     6.8   3.1290 0.0016824 **
F1:F2:trial    20    36.0     1.8
F1:id:trial    30    52.7     1.8
F2:id:trial    30    90.2     3.0
F1:F2:id:trial 60   130.0     2.2
Residuals       0     0.0
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.922 F(35,180): 60.642 AIC: 853.444 BIC: 978.329
        AdjR2    RSS df    AIC    BIC    c0                                Model AIC BIC
RLS      0.94 1230.3  3  996.8 1010.3 -1.01 0.37 0.64  NaN  NaN  NaN  NaN a+b=1
w A=B    0.89 1709.0  3 1067.7 1081.2  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.92 1241.6  3  998.7 1012.2 -0.48 1.00 1.00 1.00 2.00 2.00 2.00
w x6     0.94  981.9  7  956.1  983.1 -0.48 1.00 3.80 4.25 6.12 9.94 2.32       AIC BIC
AVERAGE  0.92 1181.0  6  993.9 1017.6 -0.54 1.00 2.04 2.04 2.20 2.04 2.29
conj.    0.88 1601.8  7 1061.8 1088.8 -0.50 1.55 2.30 1.00 1.55 2.30 1.00
AxB      0.90 2028.5  2 1102.8 1112.9  3.87 1.03  NaN  NaN  NaN  NaN  NaN

[regr] Additive model
Estimation mode: RLS

                     Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1            (Intercept)   -1.011    0.306  -1.614  -0.408  -3.305 0.0011 **
2 object@xPredictability    0.374    0.034   0.307   0.441  11.017 0.0000 ***
3    object@xBenevolence    0.643    0.024   0.596   0.689  27.235 0.0000 ***

     SSreg    SSres   Rsquare        F dfk dfres       p      AIC      BIC
1 20592.652 1230.287 0.9436241 1182.825   3   212 0.00000 996.7619 1010.263

[equal] Averaging model: overal equal weighted case
Estimation mode: W

             Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1             c0    0.000      NaN     NaN     NaN     NaN    NaN NaN
2 Predictability    0.500    0.026   0.449   0.551  19.429 0.0000 ***
3    Benevolence    0.500    0.023   0.455   0.545  21.878 0.0000 ***

     SSreg    SSres  Rsquare       F dfk dfres       p      AIC      BIC
1 13369.500 1708.963 0.886662 552.837   3   212 0.00000 1067.748 1081.249

[Optim2] Averaging model: equal weighted case
Estimation mode: W

             Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1             c0   -0.479      NaN     NaN     NaN     NaN    NaN NaN
2 Predictability    1.000    0.022   0.957   1.043  45.589 0.0000 ***
3    Benevolence    2.000    0.019   1.962   2.038 102.670 0.0000 ***

     SSreg    SSres  Rsquare       F dfk dfres       p      AIC      BIC
1 14075.864 1241.602 0.918942 801.138   3   212 0.00000 998.7394 1012.240

[Optim6] Averaging model: differential weighted case
Estimation mode: W

           Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1           c0   -0.478      NaN     NaN     NaN     NaN    NaN NaN
2 High Predic.    1.000    0.020   0.961   1.039  50.778 0.0000 ***
3 med. Predic.    3.797    0.017   3.762   3.831 217.084 0.0000 ***
4 Low  Predic.    4.245    0.020   4.206   4.284 215.554 0.0000 ***
5 High Benevo.    6.120    0.017   6.086   6.155 349.956 0.0000 ***
6 med. Benevo.    9.942    0.020   9.904   9.981 504.857 0.0000 ***
7 Low  Benevo.    2.319    0.017   2.285   2.354 132.600 0.0000 ***

     SSreg   SSres   Rsquare       F dfk dfres       p      AIC     BIC
1 15564.891 981.926 0.9406577 471.013   7   208 0.00000 956.0568 983.059

[Multiplying] Multiplying model
Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    3.871    0.308   3.264   4.478  12.564 0.0000 ***
2         AxB    1.032    0.041   0.952   1.113  25.170 0.0000 ***

     SSreg    SSres   Rsquare        F dfk dfres       p      AIC      BIC
1 19316.150 2028.522 0.9049635 1014.123   2   213 0.00000 1102.775 1112.901
```

Listing E.37: Experiment 2: Model selection and estimated weights for the couple of factors *predictability* × *competence*. All the sample.

```
####################################################
3  Predictability Competence
####################################################

Analysis of Variance Table

Response: y
               Df Sum Sq Mean Sq  F value    Pr(>F)
F1              2 1147.9   573.9 160.8368 2.491e-08 ***
F2              2 3707.3  1853.7 479.5936 1.169e-10 ***
id              3  130.1    43.4  11.6534 0.0003349 ***
trial           5   22.5     4.5
F1:F2           4  140.8    35.2  21.2680 5.680e-07 ***
F1:id           6   35.4     5.9   2.3802 0.0532649 .
F2:id           6  130.6    21.8   9.1577 1.010e-05 ***
F1:trial       10   35.7     3.6
F2:trial       10   38.7     3.9
id:trial       15   55.8     3.7
F1:F2:id       12  127.2    10.6   6.3070 4.677e-07 ***
F1:F2:trial    20   33.1     1.7
F1:id:trial    30   74.3     2.5
F2:id:trial    30   71.3     2.4
F1:F2:id:trial 60  100.8     1.7
Residuals       0    0.0
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.926 F(35,180): 64.491 AIC: 836.78 BIC: 961.665
        AdjR2    RSS df    AIC    BIC    c0                                 Model AIC BIC
RLS      0.95 1101.4  3  972.9  986.4 -2.44 0.47 0.68  NaN  NaN  NaN  NaN a+b=1
w A=B    0.89 1561.1  3 1048.2 1061.7  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.91 1163.4  3  984.7  998.2 -0.51 1.00 1.00 1.00 1.50 1.50 1.50
w x6     0.91 1054.0  5  967.3  987.6 -0.29 1.00 1.82 1.00 1.50 1.50 3.69
AVERAGE  0.93 1036.6  5  963.8  984.0 -0.53 1.00 1.89 1.89 2.28 1.89 1.89       AIC BIC
conj.    0.90 1401.5  7 1032.9 1059.9 -0.52 1.55 2.30 1.00 1.55 2.30 1.00
AxB      0.92 1766.9  2 1072.9 1083.1  3.37 1.12  NaN  NaN  NaN  NaN  NaN

[regr] Additive model
Estimation mode: RLS

                     Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1            (Intercept)   -2.445    0.289  -3.014  -1.875  -8.460 0.0000 ***
2 object@xPredictability    0.472    0.030   0.412   0.532  15.546 0.0000 ***
3     object@xCompetence    0.680    0.030   0.621   0.738  22.827 0.0000 ***

      SSreg    SSres   Rsquare        F dfk dfres       p     AIC    BIC
1 20737.654 1101.404 0.9495672 1330.539   3   212 0.00000 972.859 986.36

[equal] Averaging model: overal equal weighted case
Estimation mode: W

             Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1             c0    0.000      NaN     NaN     NaN     NaN    NaN NaN
2 Predictability    0.500    0.026   0.449   0.551  19.374 0.0000 ***
3     Competence    0.500    0.023   0.454   0.546  21.551 0.0000 ***

      SSreg    SSres  Rsquare       F dfk dfres       p      AIC      BIC
1 13189.500 1561.150 0.894164 597.033   3   212 0.00000 1048.208 1061.709

[Optim2] Averaging model: equal weighted case
Estimation mode: W

             Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1             c0   -0.511      NaN     NaN     NaN     NaN    NaN NaN
2 Predictability    1.000    0.022   0.956   1.044  44.886 0.0000 ***
3     Competence    1.500    0.020   1.461   1.539  74.896 0.0000 ***

      SSreg    SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 12291.649 1163.372 0.9135362 746.631   3   212 0.00000 984.6821 998.1832

[Optim6] Averaging model: differential weighted case
Estimation mode: W

           Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1           c0   -0.294      NaN     NaN     NaN     NaN    NaN NaN
2 High Predic.    1.000    0.021   0.958   1.042  46.935 0.0000 ***
3 med. Predic.    1.817    0.019   1.780   1.855  94.887 0.0000 ***
4 Low  Predic.    1.000    0.021   0.958   1.042  46.935 0.0000 ***
5 High Compet.    1.500    0.019   1.462   1.538  78.315 0.0000 ***
6 med. Compet.    1.500    0.021   1.458   1.542  70.403 0.0000 ***
7 Low  Compet.    3.688    0.019   3.651   3.726 192.567 0.0000 ***

      SSreg    SSres  Rsquare       F dfk dfres       p      AIC      BIC
1 11584.089 1053.962 0.916604 461.621   5   210 0.00000 967.3487 987.6004

[Multiplying] Multiplying model
Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    3.367    0.330   2.718   4.017  10.216 0.0000 ***
2         AxB    1.117    0.044   1.030   1.204  25.237 0.0000 ***

      SSreg    SSres   Rsquare        F dfk dfres       p      AIC      BIC
1 19573.688 1766.853 0.9172067 1179.837   2   213 0.00000 1072.944 1083.069
```

Listing E.38: Experiment 2: Model selection and estimated weights for the couple of factors *honesty* × *benevolence*. All the sample.

```
####################################################
4  Honesty Benevolence
####################################################

Analysis of Variance Table

Response: y
                Df  Sum Sq Mean Sq  F value     Pr(>F)
F1               2 3469.7  1734.9 217.8713 5.683e-09 ***
F2               2 2518.2  1259.1 110.0985 1.547e-07 ***
id               3  166.0    55.3  17.1625 4.099e-05 ***
trial            5   12.9     2.6
F1:F2            4  343.5    85.9  42.1204 1.799e-09 ***
F1:id            6  104.4    17.4   8.4535 2.061e-05 ***
F2:id            6   54.6     9.1   5.2684 0.0008263 ***
F1:trial        10   79.6     8.0
F2:trial        10  114.4    11.4
id:trial        15   48.4     3.2
F1:F2:id        12   59.8     5.0   2.1700 0.0249412 *
F1:F2:trial     20   40.8     2.0
F1:id:trial     30   61.7     2.1
F2:id:trial     30   51.8     1.7
F1:F2:id:trial  60  137.7     2.3
Residuals        0    0.0
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.925 F(35,180): 63.118 AIC: 887.773 BIC: 1012.659
        AdjR2    RSS df    AIC    BIC    c0                                   Model AIC BIC
RLS      0.95 1449.2  3 1032.1 1045.6 -1.58 0.66 0.46  NaN  NaN  NaN  NaN a+b=1
w A=B    0.91 1533.8  3 1044.4 1057.9  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.91 1355.9  3 1017.8 1031.3 -0.45 1.00 1.00 1.00 1.00 1.00 1.00
w x6     0.92 1053.2  6  969.2  992.8  0.08 1.00 1.86 2.71 1.00 1.00 1.83       AIC BIC
AVERAGE  0.92 1221.2  4  997.2 1014.0 -0.29 1.00 1.09 1.09 1.09 1.09 1.09
conj.    0.88 2079.7  7 1118.2 1145.2 -0.77 1.87 2.12 1.00 1.87 2.12 1.00
AxB      0.93 1590.7  2 1050.3 1060.4  4.02 0.98  NaN  NaN  NaN  NaN  NaN

[regr] Additive model
Estimation mode: RLS

                    Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1           (Intercept)   -1.576    0.188  -1.947  -1.205  -8.375 0.0000 ***
2      object@xHonesty     0.656    0.028   0.601   0.712  23.279 0.0000 ***
3 object@xBenevolence     0.458    0.028   0.402   0.514  16.066 0.0000 ***

     SSreg    SSres  Rsquare        F dfk dfres       p      AIC      BIC
1 25355.381 1449.163 0.945936 1236.424   3   212 0.00000 1032.129 1045.631

[equal] Averaging model: overal equal weighted case
Estimation mode: W

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1          c0    0.000      NaN     NaN     NaN     NaN    NaN NaN
2     Honesty    0.500    0.020   0.461   0.539  25.193 0.0000 ***
3 Benevolence    0.500    0.021   0.460   0.540  24.357 0.0000 ***

     SSreg    SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 15471.000 1533.841 0.9097997 712.775   3   212 0.00000 1044.396 1057.897

[Optim2] Averaging model: equal weighted case
Estimation mode: W

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1          c0   -0.454      NaN     NaN     NaN     NaN    NaN NaN
2     Honesty    1.000    0.019   0.963   1.037  53.590 0.0000 ***
3 Benevolence    1.000    0.019   0.962   1.038  51.812 0.0000 ***

     SSreg    SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 13664.770 1355.888 0.9097318 712.186   3   212 0.00000 1017.759 1031.260

[Optim6] Averaging model: differential weighted case
Estimation mode: W

           Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1           c0    0.080      NaN     NaN     NaN     NaN    NaN NaN
2 High Honest.    1.000    0.017   0.967   1.033  60.374 0.0000 ***
3 med. Honest.    1.856    0.017   1.823   1.890 108.355 0.0000 ***
4 Low  Honest.    2.705    0.017   2.673   2.738 163.330 0.0000 ***
5 High Benevo.    1.000    0.017   0.966   1.034  58.370 0.0000 ***
6 med. Benevo.    1.000    0.017   0.967   1.033  60.374 0.0000 ***
7 Low  Benevo.    1.827    0.017   1.793   1.860 106.615 0.0000 ***

     SSreg    SSres   Rsquare       F dfk dfres       p    AIC      BIC
1 12378.140 1053.188 0.9215872 409.397   6   209 0.00000 969.19 992.8169

[Multiplying] Multiplying model
Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    4.022    0.296   3.437   4.606  13.567 0.0000 ***
2         AxB    0.984    0.027   0.931   1.036  36.983 0.0000 ***

     SSreg    SSres   Rsquare        F dfk dfres       p      AIC      BIC
1 22194.157 1590.724 0.9331204 1485.913   2   213 0.00000 1050.261 1060.387
```

Listing E.39: Experiment 2: Model selection and estimated weights for the couple of factors *honesty* × *competence*. All the sample.

```
#####################################################
5  Honesty Competence
#####################################################

Analysis of Variance Table

Response: y
                Df  Sum Sq Mean Sq  F value    Pr(>F)
F1               2 2834.19 1417.10 157.5685 2.752e-08 ***
F2               2 2752.26 1376.13 206.7504 7.341e-09 ***
id               3   29.63    9.88   3.3571   0.04718 *
trial            5   33.60    6.72
F1:F2            4  377.10   94.28  26.2520 1.033e-07 ***
F1:id            6   53.61    8.94   2.6255   0.03630 *
F2:id            6   38.10    6.35   2.3675   0.05433 .
F1:trial        10   89.94    8.99
F2:trial        10   66.56    6.66
id:trial        15   44.14    2.94
F1:F2:id        12   41.72    3.48   1.6906   0.09170 .
F1:F2:trial     20   71.82    3.59
F1:id:trial     30  102.10    3.40
F2:id:trial     30   80.47    2.68
F1:F2:id:trial  60  123.38    2.06
Residuals        0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.909 F(35,180): 51.484 AIC: 911.937 BIC: 1036.822
        AdjR2    RSS df    AIC    BIC    c0                                Model AIC BIC
RLS      0.95 1255.7  3 1001.2 1014.7 -2.42 0.55 0.59  NaN  NaN  NaN  NaN a+b=1
w A=B    0.90 1648.9  3 1060.0 1073.5  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.91 1312.5  3 1010.7 1024.2 -0.62 1.00 1.00 1.00 1.00 1.00 1.00
w x6     0.92  981.8  5  952.0  972.3 -0.10 1.00 2.50 2.50 1.50 1.50 3.00       AIC BIC
AVERAGE  0.93 1060.7  4  966.7  983.6 -0.45 1.00 1.20 1.20 1.20 1.20 1.20
conj.    0.89 1849.9  7 1092.9 1119.9 -0.93 1.87 2.12 1.00 1.87 2.12 1.00
AxB      0.94 1386.9  2 1020.6 1030.8  3.88 0.98  NaN  NaN  NaN  NaN  NaN

[regr] Additive model
Estimation mode: RLS

                  Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1         (Intercept)   -2.419    0.245  -2.901  -1.938  -9.895 0.0000 ***
2     object@xHonesty    0.552    0.024   0.504   0.600  22.632 0.0000 ***
3 object@xCompetence    0.587    0.027   0.532   0.641  21.385 0.0000 ***

     SSreg    SSres  Rsquare        F dfk dfres       p      AIC      BIC
1 22873.549 1255.713 0.947959 1287.235   3   212 0.00000 1001.180 1014.681

[equal] Averaging model: overal equal weighted case
Estimation mode: W

         Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1         c0    0.000      NaN     NaN     NaN     NaN    NaN NaN
2    Honesty    0.500    0.021   0.458   0.542  23.400 0.0000 ***
3 Competence    0.500    0.022   0.456   0.544  22.349 0.0000 ***

     SSreg    SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 15291.000 1648.908 0.9026613 655.321   3   212 0.00000 1060.021 1073.522

[Optim2] Averaging model: equal weighted case
Estimation mode: W

         Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1         c0   -0.624      NaN     NaN     NaN     NaN    NaN NaN
2    Honesty    1.000    0.019   0.962   1.038  52.455 0.0000 ***
3 Competence    1.000    0.020   0.961   1.039  50.100 0.0000 ***

     SSreg    SSres  Rsquare       F dfk dfres       p      AIC      BIC
1 12808.618 1312.534 0.907052 689.615   3   212 0.00000 1010.740 1024.241

[Optim6] Averaging model: differential weighted case
Estimation mode: W

           Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1           c0   -0.104      NaN     NaN     NaN     NaN    NaN NaN
2 High Honest.    1.000    0.017   0.967   1.033  60.363 0.0000 ***
3 med. Honest.    2.500    0.017   2.466   2.534 144.130 0.0000 ***
4 Low  Honest.    2.500    0.017   2.467   2.533 150.906 0.0000 ***
5 High Compet.    1.500    0.017   1.466   1.534  86.478 0.0000 ***
6 med. Compet.    1.500    0.017   1.467   1.533  90.544 0.0000 ***
7 Low  Compet.    3.000    0.017   2.966   3.034 172.956 0.0000 ***

     SSreg   SSres   Rsquare       F dfk dfres       p      AIC      BIC
1 11244.658 981.839 0.9196958 481.011   5   210 0.00000 952.0376 972.2893

[Multiplying] Multiplying model
Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    3.882    0.244   3.402   4.363  15.940 0.0000 ***
2         AxB    0.981    0.021   0.939   1.023  46.191 0.0000 ***

     SSreg    SSres  Rsquare        F dfk dfres       p      AIC      BIC
1 21606.571 1386.873 0.939684 1659.200   2   213 0.00000 1020.640 1030.765
```

Listing E.40: Experiment 2: Model selection and estimated weights for the couple of factors *benevolence* × *competence*. All the sample.

```
####################################################
6  Benevolence Competence
####################################################

Analysis of Variance Table

Response: y
                Df  Sum Sq Mean Sq  F value    Pr(>F)
F1               2 2060.22 1030.11 235.3664 3.895e-09 ***
F2               2 2590.05 1295.03 276.5156 1.767e-09 ***
id               3   90.07   30.02   3.3006 0.0494504 *
trial            5   32.89    6.58
F1:F2            4  312.69   78.17  25.7741 1.202e-07 ***
F1:id            6   20.39    3.40   1.2105 0.3282532
F2:id            6  142.86   23.81   6.0981 0.0002908 ***
F1:trial        10   43.77    4.38
F2:trial        10   46.83    4.68
id:trial        15  136.45    9.10
F1:F2:id        12   97.18    8.10   2.6523 0.0064368 **
F1:F2:trial     20   60.66    3.03
F1:id:trial     30   84.22    2.81
F2:id:trial     30  117.14    3.90
F1:F2:id:trial  60  183.20    3.05
Residuals        0    0.00
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '␣' 1

Fit Indexes for model without repeted factor
 R2: 0.883 F(35,180): 38.753 AIC: 942.538 BIC: 1067.423
        AdjR2    RSS df    AIC    BIC    c0                                  Model AIC BIC
RLS      0.94 1521.4  3 1042.6 1056.1 -1.55 0.46 0.63  NaN  NaN  NaN  NaN a+b=1
w A=B    0.89 1710.6  3 1068.0 1081.5  0.00 1.00 1.00 1.00 1.00 1.00 1.00
w x2     0.89 1531.7  3 1044.1 1057.6 -0.46 1.00 1.00 1.00 1.00 1.00 1.00
w x6     0.90 1222.2  6 1001.3 1025.0 -0.20 1.00 2.55 1.00 1.00 1.51 2.41       AIC BIC
AVERAGE  0.91 1265.6  6 1008.9 1032.5 -0.31 1.00 1.73 1.52 1.52 1.53 1.52
conj.    0.90 1502.7  7 1048.0 1075.0 -0.43 1.36 2.06 1.00 1.36 2.06 1.00
AxB      0.91 1935.0  2 1092.6 1102.7  4.33 0.93  NaN  NaN  NaN  NaN  NaN

[regr] Additive model
Estimation mode: RLS

                   Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1          (Intercept)   -1.548    0.238  -2.017  -1.080  -6.509 0.0000 ***
2 object@xBenevolence     0.457    0.025   0.407   0.507  18.095 0.0000 ***
3  object@xCompetence     0.631    0.028   0.575   0.686  22.437 0.0000 ***

     SSreg    SSres  Rsquare        F dfk dfres       p      AIC      BIC
1 22967.990 1521.366 0.9378764 1066.851   3   212 0.00000 1042.632 1056.133

[equal] Averaging model: overal equal weighted case
Estimation mode: W

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1          c0    0.000      NaN     NaN     NaN     NaN    NaN NaN
2 Benevolence    0.500    0.022   0.456   0.544  22.226 0.0000 ***
3  Competence    0.500    0.023   0.455   0.545  21.957 0.0000 ***

     SSreg    SSres  Rsquare       F dfk dfres       p      AIC      BIC
1 14769.000 1710.609 0.8961985 610.120   3   212 0.00000 1067.956 1081.457

[Optim2] Averaging model: equal weighted case
Estimation mode: W

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1          c0   -0.455      NaN     NaN     NaN     NaN    NaN NaN
2 Benevolence    1.000    0.021   0.958   1.042  46.976 0.0000 ***
3  Competence    1.000    0.022   0.958   1.042  46.407 0.0000 ***

     SSreg    SSres  Rsquare       F dfk dfres       p      AIC      BIC
1 12958.255 1531.718 0.8942911 597.836   3   212 0.00000 1044.097 1057.598

[Optim6] Averaging model: differential weighted case
Estimation mode: W

           Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1           c0   -0.203      NaN     NaN     NaN     NaN    NaN NaN
2 High Benevo.    1.000    0.019   0.962   1.038  52.215 0.0000 ***
3 med. Benevo.    2.554    0.019   2.515   2.592 131.716 0.0000 ***
4 Low  Benevo.    1.000    0.019   0.962   1.038  52.215 0.0000 ***
5 High Compet.    1.000    0.019   0.962   1.038  51.582 0.0000 ***
6 med. Compet.    1.513    0.019   1.475   1.550  78.985 0.0000 ***
7 Low  Compet.    2.414    0.019   2.375   2.452 124.502 0.0000 ***

     SSreg    SSres  Rsquare       F dfk dfres       p      AIC      BIC
1 11881.303 1222.238 0.9067246 338.613   6   209 0.00000 1001.344 1024.971

[Multiplying] Multiplying model
Estimation mode: RLS

          Liv Estimate Std.Err. lim.inf lim.sup t.value      p sig
1 (Intercept)    4.333    0.335   3.672   4.994  12.916 0.0000 ***
2         AxB    0.934    0.036   0.864   1.004  26.314 0.0000 ***

     SSreg    SSres  Rsquare        F dfk dfres       p      AIC      BIC
1 21033.380 1935.047 0.9157519 1157.623   2   213 0.00000 1092.585 1102.711
```

Figure E.1: Experiment 2: Box-plot of the observed data from the two-ways designs; dashed curves show the optimal weight and value parameters for the averaging model (2nd subject)

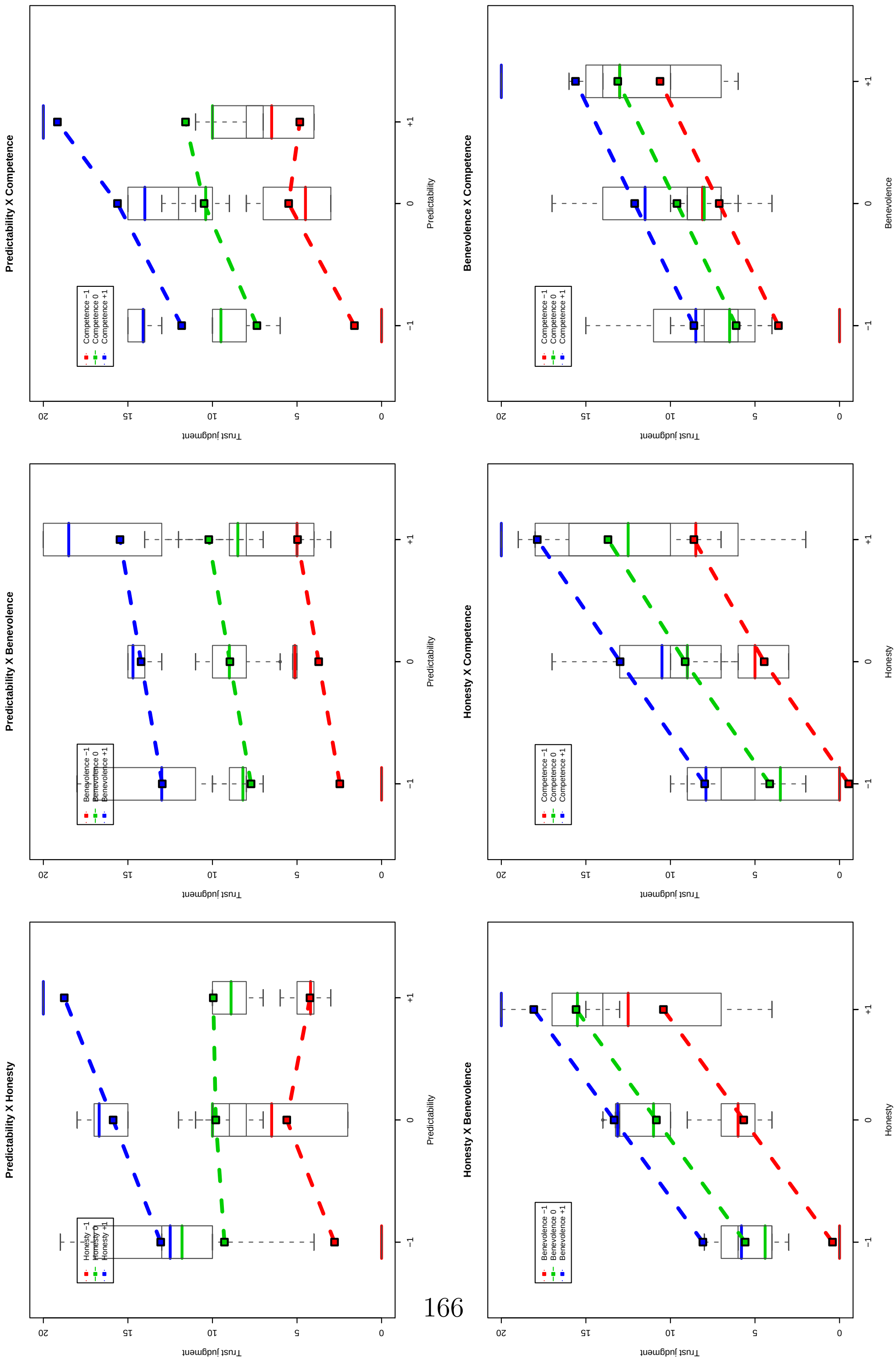


Figure E.2: Experiment 2: Box-plot of the observed data from the two-ways designs; dashed curves show the optimal weight and value parameters for the averaging model (3rd subject)

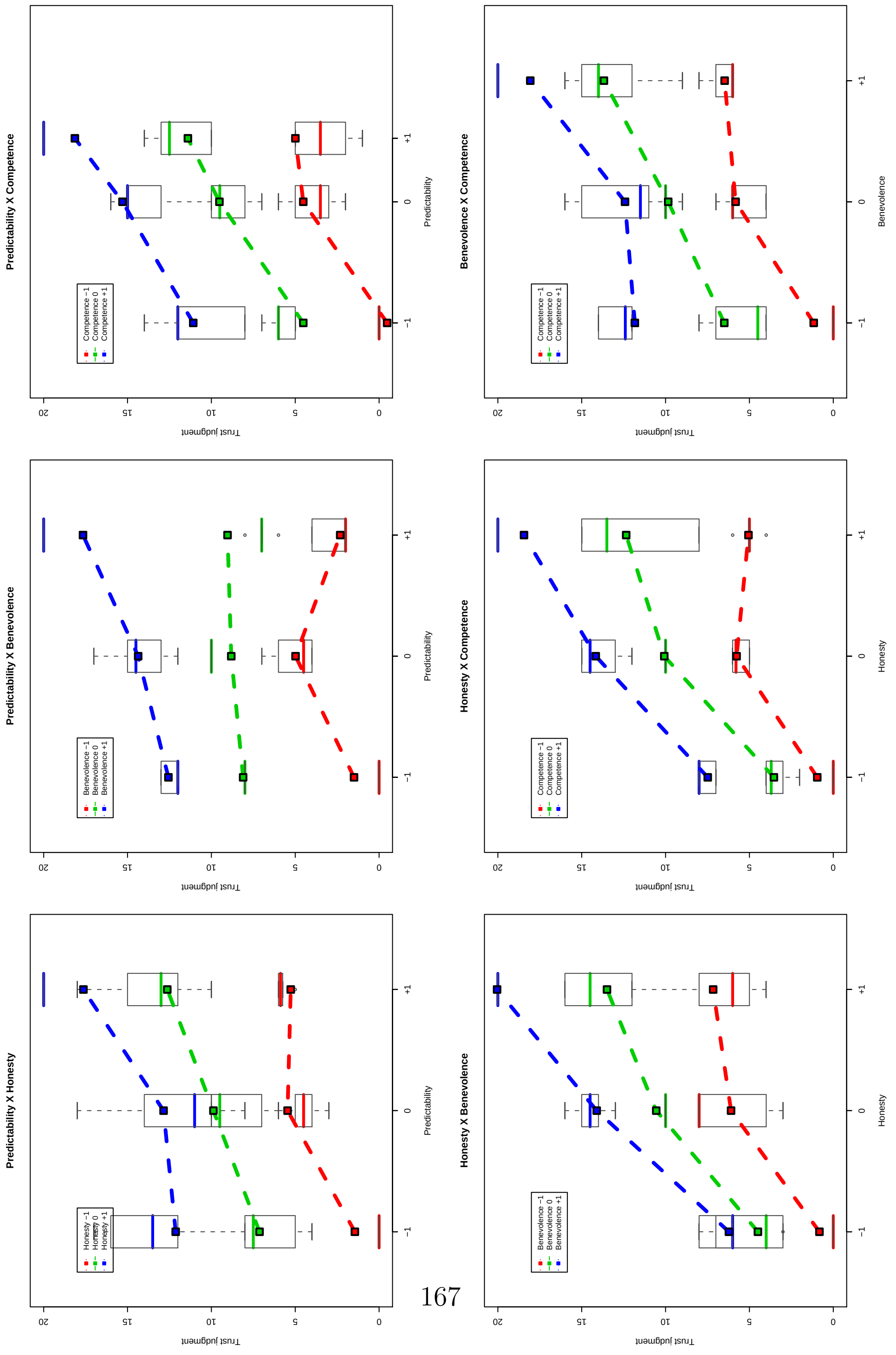
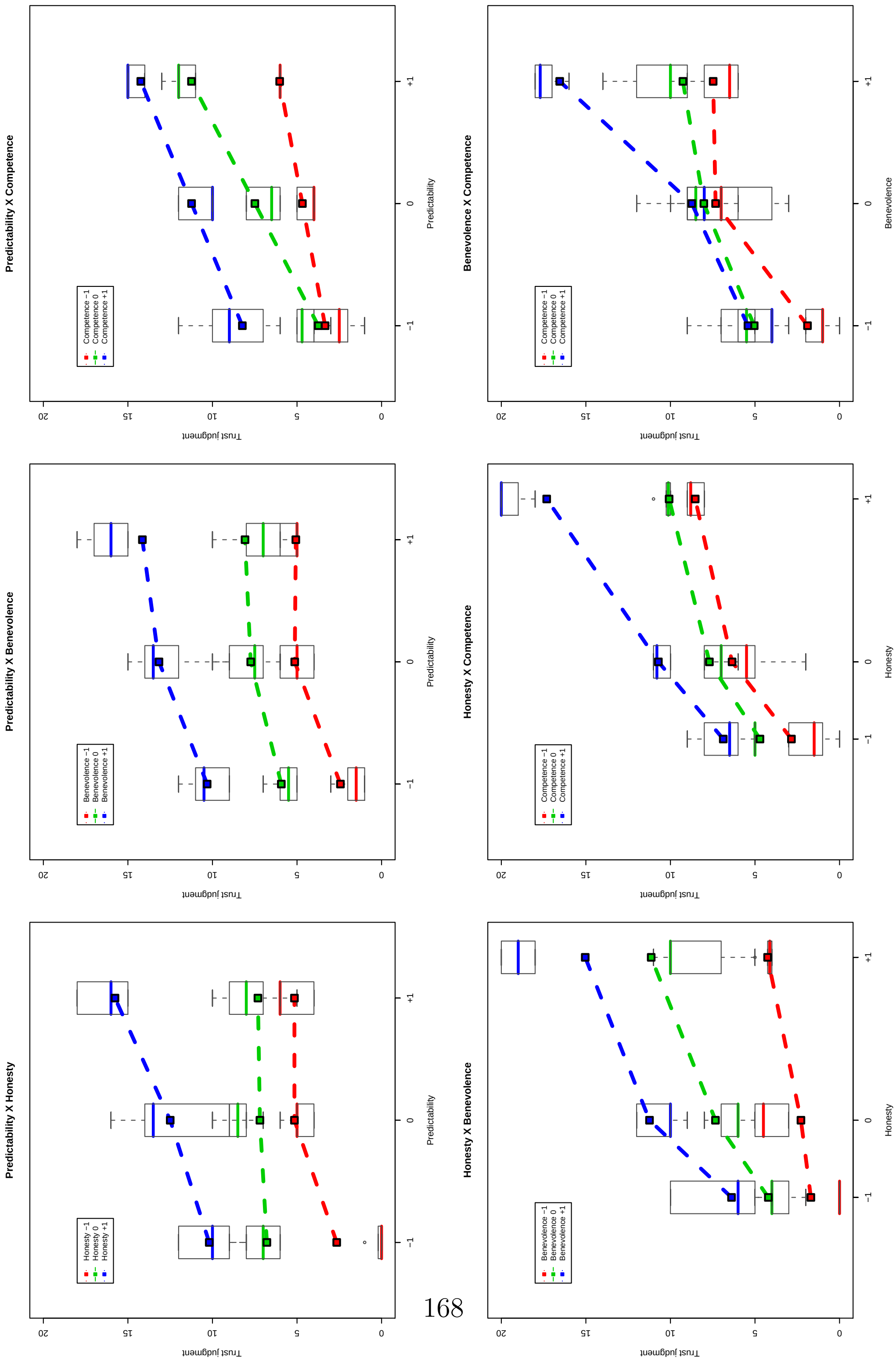


Figure E.3: Experiment 2: Box-plot of the observed data from the two-ways designs; dashed curves show the optimal weight and value parameters for the averaging model (4th subject)



Listings

List of Figures

Bibliography

Akaike, H. (1976). Canonical correlation analysis of time series and the use of an information criterion. In R. K. Mehra & D. G. Lainotis (Eds.), *System identification: Advances and case studies* (pp. 52–107). New York: Academic Press.

Anderson, N. H. (1962a). Application of an Additive Model to Impression Formation. *Science*, *138*(3542), 817–818.

Anderson, N. H. (1962b). On the quantification of Miller's conflict theory. *Psychological Review*, *69*, 410–414.

Anderson, N. H. (1965). Averaging versus adding as a stimulus combination rule in impression formation. *Journal of Experimental Psychology*, *70*, 394–400.

Anderson, N. H. (1974a). Algebraic models in perception. In E. Carterette & M. Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press.

Anderson, N. H. (1974b). Cognitive Algebra. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 7). New York: Academic Press.

Anderson, N. H. (1974c). Information Integration Theory: a brief survey. In D. Krantz, R. Atkinson, R. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2). San Francisco: Freeman.

Anderson, N. H. (1976). How functional measurement can yield validated interval scales of mental quantities. *Journal of Applied Psychology*, *61*, 677–692.

Anderson, N. H. (1980). Information Integration Theory in Developmental Psychology. In F. Wilkening, J. Becker, & T. Trabasso (Eds.), *Information Integration by Children* (pp. 1–45). Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, N. H. (1981). *Foundations of Information Integration Theory.* New York: Academic Press.

Anderson, N. H. (1982). *Methods of Information Integration Theory.* New York: Academic Press.

Anderson, N. H. (1983). Cognitive Algebra in Intuitive Physics. In H. G. Geissler (Ed.), *Modern Issues in Perception* (pp. 229–253). Amsterdam: North-Holland.

Anderson, N. H. (1989). Information integration approach to emotions and their measurement. In R. Plutchik & K. H. (Eds.), *Emotion: Theory, Research and Experience* (Vol. 4, pp. 133–186). San Diego: Academic Press.

Anderson, N. H. (1990a). Functional memory in person cognition. In N. H. Anderson (Ed.), *Contributions to Information Integration Theory* (Vol. 1: Cognition, pp. 1–55). Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, N. H. (1990b). Integration Psychophysics. In H. G. Geissler (Ed.), *Psychophysical explorations of mental structures* (pp. 71–93). Toronto: Hogrefe & Huber.

Anderson, N. H. (1992). Integration psychophysics and cognition. In D. Algom (Ed.), *Psychophysical Approaches to Cognition* (pp. 13–114). Amsterdam: North-Holland.

Anderson, N. H. (1996). *A functional theory of cognition*. Mahwah, NY: Lawrence Erlbaum Associates.

Anderson, N. H. (1997). Intuitive Physics: Understanding and Learning of Physical Relations. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: an investigation of conceptual structures and processes* (pp. 231–265). Washington, DC: APA.

Anderson, N. H. (2001a, Oct, 19th-23rd). Context effects as Foundation for Unified Theory. In *Proceedings of the 17th Annual Meeting of the International Society for Psychophysics (Fechner Day)*. Leipzig: International Society for Psychophysics.

Anderson, N. H. (2001b). *Empirical direction in design and analysis*. Mahwah, NY: Lawrence Erlbaum Associates.

Anderson, N. H. (2004). Unified Theory. In J. T. Jost, M. R. Banaji, & D. A. Prentice (Eds.), *Perspectivism in Social Psychology: the Yin and Yang of Scientific* (pp. 287–304). Washington, DC: APA.

Anderson, N. H., & Cuneo, D. O. (1978a). The Height+Width rule in Children's judgments of quantitys. *Journal of Experimental Psychology: General*, *107*(4), 335–378.

Anderson, N. H., & Cuneo, D. O. (1978b). The Height+Width rule seems solid: reply to Bogartz. *Journal of Experimental Psychology: General*, *107*(4), 388–392.

Anderson, N. H., & Zalinski, J. (1990). Functional measurement approach to self-estimation in multiattribute evaluation. In N. H. Anderson (Ed.), *Contributions to Information Integration Theory* (Vol. 1: Cognition, pp. 144–185). Hillsdale, NJ: Lawrence Erlbaum Associates.

Arens, Y., Knoblock, C. A., & Shen, W. (1996). Query reformulation for dynamic information integration. *Journal of Intelligent Information Systems*, *6*(2–3), 99–130.

Ariew, R. (1976). *Ockham's Razor: A historical and philosophical analysis of Ockham's Principle of Parsimony.* Champaign-Urbana: University of Illinois.

Aristotle. (2002). *Nicomachean ethics: Translation, introduction, commentary* (C. Rowe, Ed.). Oxford: Oxford University Press.

Bayes, T. (1763). An essay towards solving a problem in the Doctrine of Chances. *Philosophical transaction of the Royal society of London*, *53*, 370–418.

Bellio, R., & Ventura, L. (2005, October). An introduction to Robust estimation with R functions. In *Proceedings of the 1st International Workshop on Robust Statistics and* **R**. Treviso: Department of Statistics, Ca' Foscari University (Venezia).

Bozzi, P. (1990). *Fisica ingenua.* Milano: Garzanti.

Bromiley, P., & Cummings, L. L. (1995). Transactions costs in organizations with trust. In R. Bies, B. Sheppard, & R. Lewicki (Eds.), *Research on negotiations in organizations* (Vol. 5, pp. 219–247). Stanford: JAI Publishing Co.

Bubna, K., & Stewart, C. V. (2000). Model selection techniques and merging rules for range data segmentation algorithms. *Computer Vision and Image Understanding*, *80*(2), 215–245.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical-theoretic approach* (2nd ed.). New York: Springer.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods and Research*, *33*(2), 261–304.

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *Journal Scientific Computing*, *16*, 1190–1208.

Castaldo, S. (2002). *Fiducia e relazioni di mercato.* Bologna: Il Mulino.

Chandler, J. P. (1969). STEPIT finds Local Minima of a Smooth Function of Several Parameters. *Behavioral Science*, *14*, 81–82.

Charlesworth, M. J. (1956). Aristotle's Razor. *Philosophical Studies*, *6*, 105–112.

Clement, J. (1983). A Conceptual Model Discussed by Galileo and used intuitively by Physics students. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 325–340). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Mahwah, NY: Lawrence Erlbaum Associates.

Cohen, J., West, S. G., Aiken, L., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral science.* Mahwah, NY: Lawrence Erlbaum Associates.

Domingos, P. (1998, Aug 27–31). Occam's two Razors: The Sharp and the Blunt. In *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining.* New York, NY.

Edwards, W., & Newman, J. R. (1982). *Multiattribute evaluation.* Beverly Hills, CA: Sage.

Falcone, R., Pezzulo, G., & Castelfranchi, C. (2003). A Fuzzy approach to a Belief-Based Trust Computation. *Lecture Notes in Computer Science*, *2631*, 73–86.

Falconi, A., & Mullet, E. (2003). Cognitive algebra of love through the adult live. *International Journal of Aging and Human Development*, *57*(3), 275–290.

Fienber, S. E. (2003). When did Bayesian Inference become "Bayesian"? *Bayesian Analysis*, *1*(1), 1–41.

Fischer, R. (2004). Bayesian group analysis of plasma-enhanced chemical vapour deposition data. *New Journal of Physics*, *25*(6).

Fumagalli Beonio Brocchieri, M. (1996). *Storia della filosofia medievale.* Roma: Laterza.

Galileo, G. (1744). *Dialogo* (Vol. 4). Padova: Manfré.

Gambetta, D. G. (1988). Can we trust trust? In *Trust: Making and breaking cooperative relations.* New York: Basil Blackwell.

Giorello, G. (1994). *Introduzione alla filosofia della scienza.* Milano: Bompiani.

Girard, M., & Mullet, E. (1997). Propensity to forgive in adolescents, young adults, older adults, and elderly people. *Journal of Adult Development*, *4*, 209–220.

Girard, M., Mullet, E., & Callahan, S. (2002). Mathematics of forgiveness. *The American Journal of Psychology*, *115*(3), 351–375.

Girden, E. R. (1992). *ANOVA Repeated Measures* (No. 84). Newbury Park: Sage University Papers.

Gorin, J. S., Dodd, B. G., Fitzpatrick, S. J., & Shieh, Y. Y. (2005). Computerized adaptive testing with the Partial credit model: Estimation procedures, Population distributions, and Item pool characteristics. *Applied Psychological Measurement*, *29*(6), 433-456.

Grahne, G., & Kiricenko, V. (2004). Towards an algebraic theory of information integration. *Information and Computation*, *194*, 79–100.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (1986).

Robust Statistics: The approach based on influence functions. New York: Wiley.

Hendry, D. F., Lu, M., & Mizon, G. E. (2004). Model Identification and Non-unique Structure. *Journal of Econometrics*, *118*(2), 129–149.

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging. A tutorial. *Statistical Science*, *14*(4), 382–417.

Huber, P. J. (1981). *Robust Statistics*. New York: Wiley.

Jäger, S., & Wilkening, F. (2001). Development of Cognitive Averaging: When Light and Light Make Dark. *Journal of Experimental Child Psychology*, *79*, 323–345.

Jefferys, W. H. (1939). *Theory of Probability* (III ed.). Oxford: Clarendon Press.

Jefferys, W. H., & Berger, J. O. (1992). Sharpening Ockham's razor on a bayesian strop. *American Scientist*, *80*, 64–72.

Josang, A., Keser, C., & Dimitrakos, T. (2005, may). Can we manage trust? In *Proceedings of the 3rd International Conference on Trust Management (iTrust)*. Paris.

Kadane, J. B., & Lazar, N. A. (2004). Methods and Criteria for Model Selection. *Journal of the American Statistical Association*, *99*(465), 279–290.

Karpp, E. R., & Anderson, N. H. (1997). Cognitive Assessment of Function Knowledge. *Journal of Research in Science Teaching*, *34*(4), 359–376.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 730–795.

Kearney, M. D. (2002). *Classroom use of multimedia-supported predict-observe-explain tasks to elicit and promote. Discussion about students' physics conceptions*. Unpublished doctoral dissertation, Curtin University of Technology, Western Australia.

Klitzner, M. D., & Anderson, N. H. (1977). Motivation × Expectancy × Value: A functional measurement approach. *Motivation and Emotion*, *1*(4), 347–365.

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods and Research*, *33*(2), 188–229.

Kutner, M. H., Nachtsheim, C. J., Wasserman, W., & Neter, J. (2004). *Applied Linear Regression Models*. Ney York: McGraw Hill.

Leamer, E. E. (1978). *Specification searches. Ad-hoc inference with non-experimental data*. New York: Wiley.

Luna-Reyes, L. F., Cresswell, A. M., & Richardson, G. P. (2004). *Knowledge and the development of interpersonal trust: a dynamic model*. Hawaii.

Manly, B. F. J. (1997). *Randomization, bootstrap and Monte Carlo methods in Biology*. London: Chapman & Hall.

Marazzi, A. (1993). *Algorithms, Routines, and S Functions for Robust Statistics*. Pacific Grove: Wadsworth and Brooks/Cole.

Masin, S. C. (2004). Tests of Functional Measurement Theory for multiplicative models. In A. M. Oliveira, M. Teixeira, G. F. Borges, & M. J. Ferro (Eds.), *Proceedings of the 20th Annual Meeting of the International Society for Psychophysics (Fechner Day).* Coimbra: International Society for Psychophysics.

Max, L., & Onghena, P. (1999). Some issues in the statistical analysis of completely randomized and repeated measures designs. *Journal of Speech, Language, and Hearing Research*, *42*(2), 261–270.

Mayer, R. C., Davis, J. H., & Shoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review*, *20*(3), 709–734.

McCloskey, M. (1983). Naive Theories of Motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299–324). Hillsdale, NJ: Lawrence Erlbaum Associates.

McKnight, D. H., & Chervany, N. L. (1995, nov. 20–22). Trust building processes in organizational relationships. In C. T. Ragsdale, E. C. Houck, & R. T. Sumichrast (Eds.), *1995 Proceedings Decision Science Institute* (Vol. 2, pp. 751–753). Boston, MA.

McKnight, D. H., & Chervany, N. L. (2002). What Trust means in E-Commerce Customer Relationships: An interdisciplinary conceptual typology. *International Journal of Electronic Commerce*, *6*(2), 35–59.

Meneghelli, G. (2004). *Il modello Averaging nella valutazione delle credenze di fiducia.* (Unpublished thesis. University of Padua)

Murphy, P. M., & Pazzani, M. J. (1994). Exploring the decision forest: an empirical investigation of Occam's razor in decision tree induction. *Journal of Artificial Intelligence Research*, *1*.

Nelder, J. A., & Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, *7*, 308–313.

Noble, S., & Shanteau, J. C. (1999). Information Integration Theory: A Unified Cognitive Theory. *Journal of Mathematical Psychology*, *43*, 449–454.

Oden, G. C., & Anderson, N. H. (1971). Differential weighting in integration theory. *Journal of Applied Psychology*, *89*, 152–161.

Oral, M., & Kettani, O. (1989). Modelling the Process of Multiattribute Choice. *Journal of the Operational Research Society*, *40*(3), 281–291.

Ouédraogo, A., & Mullet, E. (2001). Prediction of performance among West African male farmers: Natural and supernatural factors. *International*

Journal of Psychology, *36*(1), 32–41.
Peng, J. L., & Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology*, *78*, 943–955.
Pötscher, B. M., & Srinivasan, S. (1994). A comparison of order estimation procedure for ARMA models. *Statistica Sinica*, *4*, 29–50.
Pöyhönen, M., & Hämäläinen, R. P. (2001). On the convergence of multiattribute weighting methods. *European Journal of Operational Research*, *129*(3), 569-585.
Prakasa Rao, B. L. S. (1992). *Identifiability in Stochastic Models: Characterization of Probability Distributions.* London: Academic Press.
R Development Core Team. (2005). *R: A language and environment for statistical computing.* Vienna, Austria.
Raftery, A. E. (1986). Choosing models for cross-classification. *American Sociological Review*, *51*, 145–146.
Raftery, A. E. (1993). Bayesian model selection in Structural Equation Models. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 163–180). Beverly Hills, CA: Sage.
Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.
Reilly, B. A. (1996). Self-insight, Other-insight, and their relation to Interpersonal Conflict. *Thinking and Reasoning*, *2*(2), 213–223.
Romano, D. M. (2003). *The nature of trust: Conceptual and operational clarification.* Unpublished doctoral dissertation, Louisiana State University, Baton Rouge, LA.
Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, *35*, 615–665.
Salibian-Barrera, M. (2005). *roblm: MM-regression estimators.* Vienna, Austria.
Schmidt, K. D., & Ackermann, B. (1990). The structure of internal representations and reaction-time related matching task phenomena. In *Psychophysical explorations of mental structures* (pp. 278–289). Toronto: Hogrefe & Huber.
Schönemann, P. H., Cafferty, T., & Rotton, J. (1973). A note on additive functional measurement. *Psychological Review*, *80*, 85–87.
Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.
Shanteau, J. C., & Anderson, N. H. (1969). Test of a conflict model for preference judgment. *Journal of Mathematical Psychology*, *6*(2), 312–325.
Shanteau, J. C., & Anderson, N. H. (1972). Integration Theory applied

to judgments of the value of information. *Journal of Experimental Psychology*, *92*(2), 266–275.

Singh, R. (1990). Two problems in cognitive algebra: Imputations and averaging versus multiplying. In N. H. Anderson (Ed.), *Contributions to Information Integration Theory* (Vol. 2: Social, pp. 143–180). Mahwah, NY: Lawrence Erlbaum Associates.

Singh, R. (1998). Causal analyses in interpersonal relations: On importance of the control condition. *Asian Journal of Social Psychology*, *1*(1), 33–50.

Singh, R., & Bhargava, S. (1986). Constant-weight versus relative-weight averaing in the prediction of exam performance. *Journal of Experimental Social Psychology*, *22*(6), 547–566.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1982). Facts versus fears: Understanding perceived risk. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: heuristics and biases* (pp. 463–492). Cambridge: Cambridge University press.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, *6*, 649–744.

Smith, A. F. M., & Spiegelhalter, D. J. (1980). Bayes factors and choice criteria for linear models. *Journal of the Royal Statistical Society. Ser. B*, *42*, 213–220.

Sober, E. (1982). The Principle of Parsimony. *British Journal for the Philosophy of Science*, *32*, 145–156.

Stigler, S. M. (1982). Thomas Bayes' Bayesian Inference. *Journal of the Royal Statistical Society.*, *145*, 250–258.

Stigler, S. M. (1983). Who discovered Bayes's theorem? *The American Statistician*, *37*(4), 290–296.

Thorburn, W. M. (1915). Occam's Razor. *Mind*, *24*, 287–288.

Tononi, G. (2004). An information integration theory of consciusness. *BMC Neuroscience*, *5*(42), 1–22.

Tornay, S. C. (1938). *Ockham: studies and selections*. La Salle, IL: Open court.

Tyler, T. R., & Kramer, R. M. (1996). Whither trust? In R. M. Kramer & T. R. Tyler (Eds.), *Trust in Organizations* (pp. 1–16). Thousand Oaks: Sage.

Ullman, J. D. (2000). Information integration using logical views. *Theoretical Computer Science*, *239*(2), 189–210.

Wang, M., & Yang, J. (1998). A multi-criterion experimental comparison of three multi-attribute weight measurement methods. *Journal of Multi-criteria Decision Analysis*, *7*, 340–350.

Webb, G. I. (1996). Further experimental evidence against the utility of Occam's razor. *Journal of Artificial Intelligence Research*, *4*, 397–417.

Weinfurt, K. P. (2000). Repeated measures analyses. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (chap. 10). Washington, DC: APA.

Weisberg, S. (2005). *Applied Linear Regression*. New York: Wiley.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*(8), 1293–1313.

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*(8), 1314–1329.

Wichura, M. J. (1988). Algorithm AS 241: The Percentage Points of the Normal Distribution. *Applied Statistics*, *37*, 477–484.

Wilkening, F., & Anderson, N. H. (1982). Comparison of the Two Rule-Assessment Methodologies for studying Cognitive Development and Knowledge Structure. *Psychological Bulletin*, *92*(1), 215–237.

Wilkening, F., Schwarz, G., & Rümmele, A. (1997). The developmental emergence of multiplicative combinations. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Conceptual structures and processes: Emergence, discovery, and change* (pp. 111–127). Washington, DC: APA.

Wrinch, D., & Jefferys, W. H. (1921). On certain fundamental principles of scientic inquiry. *Phil. Mag.*, *42*, 369–390.

Yamagishi, T., & Hill, C. T. (1981). Adding versus averagin models revisited: A test of a path-analytic integration model. *Journal of Personality and Social Psychology*, *41*(1), 13–25.

Yamagishi, T., & Hill, C. T. (1983). Initial impression verso missing information as explanations of the set-size effect. *Journal of Personality and Social Psychology*, *44*, 942–951.

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion*, *18*(2), 129–166.

Zalinski, J. (1987). *Parameter estimation for the averaging model of information integration theory.* Unpublished doctoral dissertation, University of California, San Diego, La Jolla, CA.

Zalinski, J., & Anderson, N. H. (1986). *Average: a user-friendly fortran-77 program for parameter estimation for the averaging model of information integration theory [computer software].* San Dieco, CA.

Zalinski, J., & Anderson, N. H. (1989). Measurement of importance in multiattribute models. In J. B. Sidowski (Ed.), *Conditioning, cognition and methodology. Contemporary issues in experimental psychology* (pp. 177–215). Lanham, MD: University presso of America.

Zalinski, J., & Anderson, N. H. (1990). Parameter estimation for averaging theory. In N. H. Anderson (Ed.), *Contributions to Information Integration Theory* (Vol. 1: Cognition, pp. 353–394). Hillsdale, NJ: Lawrence Erlbaum Associates.

Zhu, S. H., & Anderson, N. H. (1991). Self-estimation of weight parameter in multiattributable analysis. *Organizational Behavior & Human Decision Processes*, *48*(1), 36–54.

Zicari, M. (2004). *Analisi delle credenze nel modello interdisciplinare di fiducia di Mcknight e Chervany.* (Unpublished thesis. University of Padua)