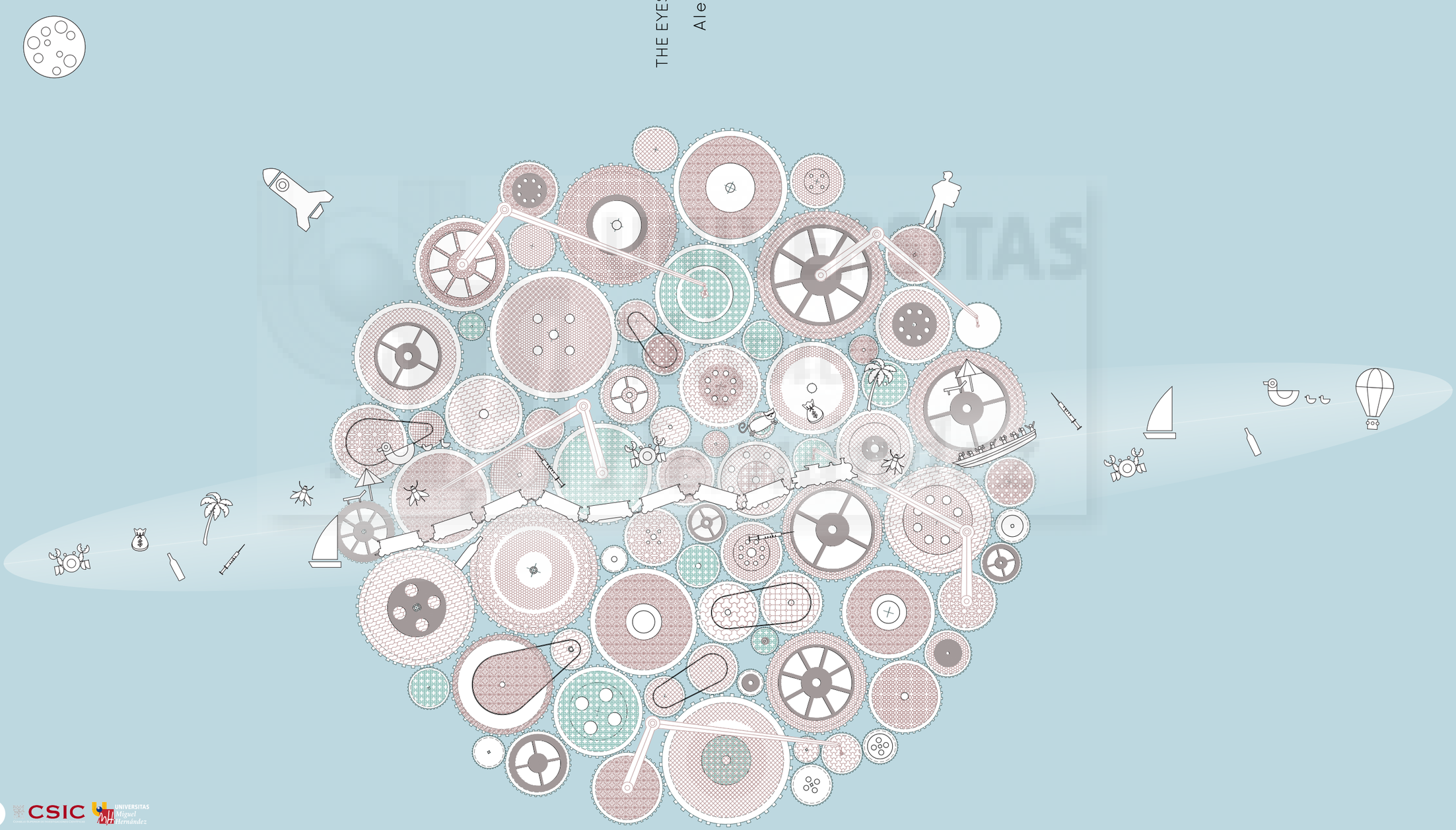
THE EYES OF THE MORAL MIND

Alexandra Gomis Pont

THE EYES OF THE MORAL MIND:
affect based-gain control and contextual modulation of moral decisions



CSIC

UNIVERSITAS Miguel Hernández

INSTITUTO DE NEUROCIENCIAS

Alexandra Gomis Pont

Supervisor: Luis M. Martinez Otero

UNIVERSITAS
Miguel
Hernández

THE EYES OF THE MORAL MIND:

Affect based-gain Control and Contextual Modulation of Moral Decisions

Presented by:

Alexandra Gomis Pont

For the degree of Doctor in Neuroscience from the Universidad Miguel Hernández

Director:

Luís M. Martínez Otero

San Juan de Alicante, 2016

UNIVERSITAS
Miguel
Hernández

UNIVERSITAS Miguel Hernández

CSIC UNIVERSITAS Miguel Hernández
INSTITUTO DE NEUROCIENCIAS

Comisión Académica del Programa
de Doctorado en Neurociencias
Universidad Miguel Hernández

Prof. Salvador Martínez
Director

A QUIEN CORRESPONDA:

Prof. Salvador Martínez Pérez, Director del Instituto de Neurociencias, centro mixto de la Universidad Miguel Hernández, UMH y la Agencia Estatal Consejo Superior de Investigaciones Científicas, CSIC

CERTIFICA:

Que la Tesis Doctoral titulada: "Neuronal *" The Eyes of the Moral Mind: Affect based-gain Control and Contextual"* ha sido realizada por D.ª Alexandra Gomis Pont (NIF 48571681M) bajo la dirección del Dr. Luís Miguel Martínez Otero y da su conformidad para que sea presentada a la Comisión de Doctorado de la Universidad Miguel Hernández.

Para que así conste a los efectos oportunos, firma el presente certificado en San Juan de Alicante a 16 de septiembre de 2016

INSTITUTO DE NEUROCIENCIAS

Salvador Martínez
Director

direccion.in@umhc.es
director.in@umh.es
www.ina.umh.es

Tel: +34 965-919221(d)
+34 965-919220 (s)
Fax: +34 965 919561

Av Ramón y Cajal s/n
Campus de San Juan
03550 SAN JUAN DE ALICANTE– ESPAÑA

UNIVERSITAS
Miguel
Hernández

Comisión Académica del Programa
de Doctorado en Neurociencias
Universidad Miguel Hernández

Prof. Salvador Martínez
Director

A QUIEN CORRESPONDA:

Prof. Salvador Martínez Pérez, Director del Instituto de Neurociencias, centro mixto de la Universidad Miguel Hernández, UMH, y la Agencia Estatal Consejo Superior de Investigaciones Científicas, CSIC

INFORMA,

Que en relación con la presentación de la Tesis Doctoral titulada " *The Eyes of the Moral Mind: Affect based-gain Control and Contextual*" de Alexandra Gomis Pont (NIF 48571681M) para la obtención del título de Doctorado Internacional y a la vista de la documentación necesaria aportada:

- La doctorando realizó una estancia en el Centre for Systems Neuroscience de la Universidad de Leicester, U.K. en el Laboratorio del Profesor Rodrigo Quian Quiroga, entre el 28 de abril y el 22 de agosto de 2014. Durante la misma, desarrolló trabajos de investigadora en formación relacionados con su tesis doctoral, reconocemos dicha estancia suficiente para cumplir la normativa correspondiente en su artículo 1.1.

En Sant Joan d'Alacant, 16 de septiembre de 2016

Salvador Martínez
Director

INSTITUTO DE NEUROCIENCIAS
CSIC
UNIVERSITAS Miguel Hernández

direccion.in@umhc.es
director.in@umh.es
www.ina.umh.es

Tel: +34 965-919221(d)
+34 965-919220 (s)
Fax: +34 965 919561

Av Ramón y Cajal s/n
Campus de San Juan
03550 SAN JUAN DE ALICANTE– ESPAÑA

UNIVERSITAS
Miguel
Hernández

INSTITUTO DE NEUROCIENCIAS

Consejo superior de Investigaciones Científicas
Universidad Miguel Hernández

A QUIEN CORRESPONDA:

El Dr. LUÍS M. MARTÍNEZ OTERO, Científico Titular del Consejo Superior de Investigaciones Científicas (CSIC) en el Instituto de Neurociencias de Alicante, Centro Mixto de CSIC-Universidad Miguel Hernández,

CERTIFICA,

Que Doña Alexandra Gomis Pont, Máster en Neurociencias, ha realizado bajo su Dirección el trabajo experimental que recoge su Tesis Doctoral "THE EYES OF THE MORAL MIND: AFFECT BASED-GAIN CONTROL AND CONTEXTUAL MODULATION OF MORAL DECISIONS". Que ha revisado los contenidos científicos y los aspectos formales del trabajo y da su conformidad para su presentación y defensa pública.

Y para que así conste a los efectos oportunos, firma el presente Certificado en San Juan de Alicante, a 26 de Septiembre de 2016.

Fdo.: Luís M. Martínez Otero

UNIVERSITAS
Miguel
Hernández

A Vito

A mi familia

Gracias por todo.

UNIVERSITAS
Miguel
Hernández

"... incluso cuando la mente opera por sí sola y, experimentando el sentimiento de condena o aprobación, declara un objeto deforme y odioso, otro bello y deseable, incluso en ese caso, sostengo que esas cualidades no están realmente en los objetos, sino que pertenecen totalmente al sentimiento de la mente que condena o alaba."

David Hume

UNIVERSITAS
Miguel
Hernández

Acknowledgements

Summary

A central aim of current research in moral decision making is to propose a mechanism to explain how a moral choice is made. Most current studies are based on the traditional postulates of the Dual Process Model (DPM) (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Haidt, 2001; Cushman, Young, & Hauser, 2006; Greene, 2007; Bartels, 2008). The automatic emotional process is posited to underlie the aversion to doing harm in up close personal dilemmas, refusing the behavior as morally acceptable, while the controlled cognitive process operates to promote welfare, maximizing choices that opt for greater number of lives saved. However, the simple division between emotional and rational components might not provide a complete picture of the moral mind. This point is appreciated by both critics and supporters of DPM arguing that affective components must be integrated in the process (Moll, Oliveira-Souza, & Zahn, 2008; Cushman, Young, & Greene, 2010; Kvaran, & Sanfey, 2010).

We propose that moral decisions are shaped by the same gain control mechanism that operates in sensory pathways representing values in a relative rather than absolute manner (Carandini, & Hegger, 2012). Context should thus strongly influence how people make affective perceptions about actions and actors in a very explicit way that is independent of purely deontological (normative ethical position) or absolute utilitarian considerations (best moral action is the one that maximizes utility).

To test this hypothesis we have designed 15 new, ecological moral dilemmas in which both the actors and actions were kept constant while context was sequentially altered as more information was added to the moral scenes. We have used these dilemmas in four different experiments where participants were asked about three aspects of the choice. Firstly, we were interested in the moral acceptability (Acc) of the dilemmas. Secondly, we also asked participants to evaluate the Stereotype Affect (SA), a basic reaction about the likeability of the active character on the dilemma. And finally, we have recorded the Perceived Affect (PA). This variable is the same as the previous one, but in this case, participants should judge the protagonist when she/he was involved in a given situational context.

Our results show that subjects judge the protagonist of the moral action encoding a subjective value (PA), which is directly dependent on the contextual value of the scenes

embedding the actions. This subjective value correlates linearly with moral acceptability. This context-dependent modulation of moral decision making is precisely described by divisive normalization, an adaptive form of gain control that may be a general mechanism for sensory and cognitive computations providing an explanation for several otherwise puzzling phenomena about decision making in general (Carandini, & Hegger, 2012).

To test the dynamics of moral judgments, we have conducted an experiment where 35 subjects made moral decisions while their eye scan paths and computer mouse trajectories were continuously monitored to measure Reaction Times (RTs), response type and uncertainty creating a Doubt Index (DI).

We have found that cognitive load correlates with behavioral responses. Low cognitive load mainly occurs in trials that corroborate the acceptance of a previously evaluated moral scene, while high cognitive load mainly characterize trials where the responses were deontological or the first contextual (the first accepted). Regarding to the RT's recordings, we have not found statistically significant differences between contextual and deontological responses.

Both results, the cognitive load pattern and the RT, are not consistent with the proposal of the DPM, where the deontological trials resulted of the fast automatic process, while the utilitarian ones from the slower and more rational and controlled process.

In addition, we have devised the Doubt Index (DI) with two different aims. First, just considering a typical behavioral DI pattern, obtained from the eye scan paths recorded during the last two trials of each dilemma, we were able to predict the participant´s responses above chance. Second, using signal detection theory to compare objective and subjective awareness of cognitive load, we were able to show that 63.2% of the participants were aware of their own deliberative process.

Finally, we have applied a Rasch Model to categorize the dilemmas depending on their difficulty and the participants' ability (Baron, Gürçay, Moore, & Starcke, 2012). We are able to predict the participant's output with these two criteria.

In summary, a more precise characterization than the classical DPM within the moral domain is based on the value representation assigned to the moral scene by divisive normalization. Its explanatory power becomes especially strong when a single computation is enough to explain all the dilemmas presented. Current research in computational neuroscience can improve this new approach to highlight one of the most interesting domains for cognitive researchers: morality.

Resumen

Uno de los objetivos principales en lo que concierne a la investigación acerca de toma de decisiones morales es la propuesta del mecanismo responsable de llevar a cabo este proceso. La mayoría de los estudios se basan en los postulados tradicionales que provienen de los Modelos Duales de Procesamiento (DPM) (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Haidt, 2001; Cushman, Young, & Hauser, 2006; Greene, 2007; Bartels, 2008). Los DPM proponen que las personas realizamos dos operaciones mentales diferentes dependiendo del tipo de información recibida. En el caso de los dilemas personales, se produce un tipo de procesamiento más automático y emocional, cuya respuesta es la de no aceptar la conducta como moral, mientras que en los dilemas impersonales, se emplea un tipo de procesamiento más racional que tiene como objetivo promover el bienestar y maximizar las decisiones que optan por salvar a un número mayor de vidas. Sin embargo, esta simple división entre componentes emocionales y racionales parece ser insuficiente. En este punto están de acuerdo tanto los críticos como los que apoyan el DPM argumentando que los componentes afectivos deberían estar integrados en este proceso (Moll, Oliveira-Souza, & Zahn, 2008; Cushman, Young, & Greene, 2010; Kvaran, & Sanfey, 2010).

Nuestra propuesta va más allá y pone al mismo nivel los estímulos morales, los dilemas, y estímulos sensoriales como puede ser la luminancia. Pensamos que responsable de la toma de decisiones morales es el mismo mecanismo de control de ganancias que opera en procesamientos sensoriales, en los cuales los valores están representados en términos relativos y no absolutos (Carandini, & Hegger, 2012). De esta manera, el contexto es el principal responsable de influenciar cómo las personas perciben el afecto en las diferentes escenas morales independientemente de motivaciones puramente deontológicas (posiciones éticas normativas) o consideraciones utilitaristas (la mejor acción moral es la que maximiza la utilidad).

Para probar esta hipótesis hemos diseñado 15 nuevos y ecológicos dilemas en el que tanto los actores como las acciones se mantenían constantes mientras que el contexto se alteraba a medida que se le iba añadiendo más información en cada una de las escenas. Hemos utilizado estos dilemas en cuatro experimentos diferentes donde se preguntaba a los participantes por tres cuestiones. En primer lugar, estábamos interesados en la aceptabilidad

(Acc) moral de los dilemas. En segundo lugar, preguntamos a los participantes por el Afecto del Estereotipo (SA); se trata de una reacción acerca de cuánto de bien cae el protagonista del dilema. Y finalmente, hemos registrado el Afecto Percibido (PA), esta variable es la misma que la anterior, pero en este caso, los participantes debían juzgar al protagonista en ese contexto.

Nuestros resultados muestran que los sujetos juzgan al protagonista de una acción moral codificando un valor subjetivo (PA), el cual es directamente dependiente del valor contextual de las escenas. Este valor subjetivo correlaciona linealmente con la aceptabilidad moral (r_s=0.9396; n=40 trials). Esta modulación contextual dependiente de la decisión moral está precisamente descrita por la normalización divisiva, una forma adaptativa de control de ganancia que podría actuar como un mecanismo general para las computaciones sensoriales y cognitivas (Carandini, & Hegger, 2012).

Para evaluar la dinámica de los juicios morales, hemos llevado a cabo un experimento en el que 35 sujetos tomaban decisiones morales mientras se registraban los movimientos oculares, las trayectorias del mouse para monitorizar los Tiempos de Reacción (RT), el tipo de respuesta y el Índice de Duda (DI).

Hemos encontrado que la carga cognitiva correlaciona con las diferentes conductas. La baja carga cognitiva ocurre principalmente en los trials que corroboran una decisión de aceptabilidad de una escena moral previa, mientras que la alta carga cognitiva ocurre en aquellos trials con respuesta deontológica o primera respuesta contextual. En relación a los registro de los RT, no hemos encontrado diferencias estadísticas entre las respuestas deontológicas y contextuales.

Ambos resultados, el patrón de carga cognitiva y los RT, no son consistentes con la propuesta del DPM, en el que las respuestas deontológicas resultan de un procesamiento automático, mientras que las utilitaristas utilizan un tipo de procesamiento más racional y controlado. Además hemos diseñado un Índice de Duda (DI) con dos objetivos diferentes. Primero, somos capaces de predecir la respuesta del sujeto por encima del nivel de azar, considerando un prototipo de DI de un patrón de respuesta, obtenido de los registros de los movimientos oculares durante los dos últimos trials. Segundo, utilizando la teoría de detección de señales hemos comparado el nivel de consciencia objetiva y subjetiva de la carga cognitiva, somos capaces de mostrar que el 63.2% de los participantes son conscientes de su propio proceso deliberativo.

Finalmente, hemos aplicado un Rasch Model para categorizer los dilemas dependiendo de su dificultad y de la habilidad del sujeto (Baron, Gürçay, Moore, & Starcke, 2012). Somos capaces de predecir el output del participante con esos dos criterios.

En conclusión, proponemos una caracterización más precisa basada en el valor de la representación asignada a la escena moral por la normalización divisiva. Su potencia explicativa se hace evidente cuando una misma computación es suficiente para explicar todos los dilemas presentados. La actual investigación en neurociencia computacional puede mejorar esta nueva aproximación para resaltar uno de los dominios más interesantes para los investigadores cognitivos: la moral.

Index

List of Figures

List of tables

Abbreviations

α Cronbach's Alpha

Acc Acceptability

AI Acceptability Index

BL Background Luminance

BSI Brief Symptom Inventory

C1 Contextual Cluster. First Contextual Response in trial 1

C2 Contextual Cluster. First Contextual Response in trial 2

C3 Contextual Cluster. First Contextual Response in trial 3

C4 Contextual Cluster. First Contextual Response in trial 4

CA Contextual Affect

CL Center Luminance

CoR Confidence Response

CR Contextual Response

D Deontological Cluster

d' Sensitivity Index

DDM Drift Diffusion Model

DI Doubt Index

DifI Difficult Index

DN Divisive Normalization

DPM Dual Process Model

DR Deontological Response

DV Dependent Variable

FN False Negative

FNR False Negative Rate

FP False Positive

FPR False Positive Rate

GLM_z Generalized Linear Models

ICC Item Characteristic Curve

IV Independent Variable

K Kruskal-Wallis statistic test

LCA Leaky Competing Accumulator Model

LSB-50 Listado de Síntomas Breves

PA Perceived Affect

PL Perceived Luminance

RT Reaction Time

SA Stereotype Affect

SCM Stereotype Content Model

SDT Signal Detection Theory

SIM Social Intuitionist Model

TAFC Two Alternative Force Choice

TN True Negative

TNR True Negative Rate

TP True Positive

TPR True Positive Rate

r_s Spearman Correlation

r Pearson Correlation

U Mann-Whitney statistic test

χ^2 Chi-square

UNIVERSITAS
Miguel
Hernández

Introduction

Moral issues greet us each morning in the newspaper, confront us in soccer fields, and bid us good night on the evening news. We are bombarded daily with questions about the justice of our foreign policy, the morality of medical technologies, the rights of the immigrants, the fairness of children, etc. Dealing with these moral issues is often perplexing. What is the right option? What factors should we consider?

Morality deals with behavioral repertoire comprising the standards of right or wrong choices. This topic became a milestone in Philosophy and over the centuries occupied hundreds of philosophical essays. The word "moral" immediately carries the concept of standards, but these standards are not always shared. Moreover, there is a low congruency between dilemmas' acceptability, showing the difficulty related with the idea of the existence of a consensual moral theory.

This topic has become a complicated issue in the multi-cultural world we live in today and a very popular topic to be studied (Greene, 2015). But one idea seems to be clear and shared by the community: "*morality is in large part about binding people together*" (Haidt, 2001). Émile Durkheim attempted to answer the question of what holds the society together, deciding that moral rules are the key aspects for the survival of any human society (Durkheim, 1973).

This introductory section provides a brief overview of which were the most important moral theories in the past, how these postulates influenced the current dominant perspective and the approach that is probably overshadowing the prevailing idea. The Appendixes 1 to 4 provide additional information about some of the issues discussed here. Finally, we present a new mechanistic explanation of the neuronal computations underlying a moral decision.

1. Main Moral Philosophical Approaches

1.1. The Philosophy of Morality: Normative Ethics

Human moral cognition has remained largely in the domain of traditional philosophy. For centuries, philosophers, including Plato, Aristotle, Immanuel Kant, David Hume, John Stuart Mill and Jeremy Bentham, created the basis of the current moral disciplines (Cushman, Young, & Greene, 2010).

1.1.1 Virtue Ethics

Virtue ethics' founding fathers are Plato and, more particularly Aristotle. This school of thought emphasizes the role of one's character and the virtues that one's character embodies for determining or evaluating ethical behavior. The concept of a virtue is the concept of something that makes its possessor good: a virtuous person is a morally good, excellent or admirable person who acts and feels well, rightly, as she should (Statman, 1997).

1.1.2 The Deontology Approach

Morality has traditionally been regarded as a code of values guiding the choices that determine the purpose and the course of our actions. The most traditional moral theory rests on principles that determine whether an action is right or wrong. Immanuel Kant produced the most influential deontological theory in the late 18th century (Kant, 1785/1964).

Kant developed his moral philosophy focusing on the maximum of the action derived from the dictates of pure reason and the categorical imperative. Reason, separate from all empirical experience, can determine the principle according to which all ends can be determined as moral. It is this fundamental principle of moral reason that is known as the categorical imperative. Pure practical reason is the process of determining what ought to be done without reference to empirical contingent factors. In Kant's view, a person cannot decide whether a certain behavior is "right," or moral, through empirical means. Such judgments must be reached *a priori*, using pure practical reason (Kant, 1785/1964).

Moral questions are determined independently of reference to the particular subject posing them. This is because morality is determined by pure practical reason rather than particular empirical or sensuous factors that morality is universally valid. This moral universalism has come to be seen as the distinctive aspect of Kant's moral philosophy and has had wide social impact in the legal and political concepts of human rights and equality.

1.1.3 Moral is based on Emotions

In the same period, a fundamental new approach was presented by David Hume (Hume, 1739/1978) who deserves credit as the first secular modern philosopher, who proposed the idea that people base moral judgments on emotions.

Hume, contrary to Kant, claimed that moral decisions are not derived from reason but rather from sentiment. While moral rationalists tend to say, first, that moral qualities are discovered by reason, and also that what is morally right is in accord with reason and what is morally evil is unreasonable, Hume rejects both principles.

Hume's position in ethics is based on four tenets:

- Reason alone cannot be the driving force of our behavior, but rather is the "slave of the passions".
- Moral decisions are not derived from reason.
- Moral decisions are derived from moral sentiments: feelings of approval (esteem, praise) and disapproval (blame) felt by spectators who contemplate a character trait or action.
- While some virtues and vices are natural, others, including justice, are artificial.

Unfortunately for him, at the time, Hume's major work on morality was rather unpopular, and his influence, at that moment, remained limited.

1.1.4 The Utilitarian Approach

Utilitarianism was conceived in the 19th century by Jeremy Bentham and John Stuart Mill to help legislators determine which laws were morally best (Mill, 1861/1998). Both Bentham and Mill suggested that ethical actions are those that provide the greatest balance of good over evil.

Utilitarianism is a version of consequentialism that results from combining consequentialism with welfarism. Consequentialism holds that an act is right if and only if it leads to the best consequences, and welfarism holds that goodness of an outcome is ultimately a matter of the amount of individual well-being, counting everyone equally; it follows that utilitarianism is the view that an act is right if and only if it leads to the greatest total amount of well-being (Kahane, 2013).

1.1.5 Normative Ethics Summary

The difference between these fours perspectives tends to concentrate more on the way in which moral dilemmas are approached, rather than on the moral conclusions reached. For example, a virtuous person or a deontologist one might argue that stealing is always wrong since this action goes against a good person and also against a universal law. A humean version would perceive stealing as wrong because it feels bad. A utilitarian perspective may argue that stealing is wrong because of the negative consequences produced by taking someone's savings without right or permission.

But despite their differences, almost all of them have much in common, including an emphasis on parsimony (ethics can be derived from a single rule), the insistence that moral decisions must be reasoned; and their focus on the abstract and universal, rather than the concrete and particular.

2. The Moral Psychology: the Rationalist Period

For decades, the conscious-reasoning perspective has been central for research in moral decision making. Yet, by the 1970s, during the cognitive period of Psychology, rational theories governed new advances in Psychology and also in moral research. Jean Piaget and Lawrence Kohlberg focusing on problems of justice, defined the characteristics of a morally mature person, and attempted to explain how experience guides a child from moral immaturity to maturity. Piaget formulated three stages of moral development (Piaget, 1932) and Lawrence Kohlberg (Kohlberg, 1958, 1969, 1971, 1984) followed Piaget's approach for studying children's conceptions of morality. Kohlberg assessed morality by asking children to consider certain situations in which right and wrong actions are not always clear.

He was not concerned with whether the children decided that certain actions were right or wrong, but with their reasoning (the way they arrived at their conclusions). The story of "Heinz Steals the Drug" is one of his best known examples (Kohlberg, 1958). The main aspect of Kohlberg's approach was the role of reason in a moral judgment. He considered moral reasoning to be a result of cognitive processes that may exist even in the absence of any kind of emotions. Based on his study of children's responses to such dilemmas, Kohlberg expanded Piaget's three stages into six, organized into three levels, each level consisting of two stages (Kohlberg, 1958, 1983). See Appendix 1 for a more detailed explanation.

Although Kohlberg's detailed investigation was a milestone in the psychological study of morality, some inconsistencies were found. For example, a person who justified a decision on the basis of principled reasoning in one situation (post conventional morality, stage 5 or 6) would frequently fall back on conventional reasoning (stage 3 or 4) in another story. Furthermore, individuals do not always progress linearly through the different stages finding, for instance, that one in fourteen people actually slipped backwards (Rest, 1980).

Piaget and Kohlberg nonetheless deserve credit for leading the field of moral Psychology. Both inspired a lot of researchers and Kohlberg's proposal dominated the field for fifteen years. But in the 1980s, research interests shifted towards the role of emotion in many psychological constructs; the "affective revolution" was coming.

3. The Dual Process Model in Moral Judgments

One of the most important contributions from the last century that has led to a better understanding of decision making is the idea of bounded rationality: when individuals make decisions, their rationality is limited by the available information, the tractability of the decision problem, the cognitive limitations of their minds, and the time available to make the decision (Simon, 1967; Kahneman, & Tversky, 1979, 1981; Kahneman, 2011). This perspective overturns the notion of human irrationality in cognitive process and has since become widely accepted. In this sense, dual systems frameworks for understanding human judgments, reasoning and behavior have been extensively influential (Epstein, 1994; Sloman, 1996; Stanovich, & West, 2000; Kahneman, 2011). Dual Process Models (DPM) are the current dominant class of theories of human decision making. They argue for the existence of two

separate, often opposing decision systems: one operates generally automatically and quickly, with little or no effort and no sense of voluntary control; whereas the other is slow, deliberative or explicit, often associated with the effortful mental activities that demand complex computations (Kahneman, 2003). Choice ultimately results from the interaction between the two.

DPM approaches captured general attention with the publication of two books (see Apendix II) by Antonio Damasio and Frans de Waal (Damasio, 1994; de Waal, 1996) and their relevance for moral decision making was further emphasized by the work of Joshua Greene (Greene, et al., 2001). Current research on human morality, where one of the most important authors is Greene, has largely focused on hypothetical scenes with non-essential contextual information (Foot, 1967; Thomson, 1985). Probably the most famous ones are the Trolley paradigm first introduced by Philippa Foot and Judith Thomson's Footbridge problem (Foot, 1967; Tomson, 1985) (see BOX 1 & 2). Both philosophical problems have inspired a large body of productive research offering diverse proposals for understanding why almost everybody responds NO to the Footbridge problem and YES to the Trolley one (Petrinovich, O'Neill, & Jorgensen, 1993). For a more detailed explanation see Appendix III.

These hypothetical dilemmas have been widely used, and they come with both advantages and disadvantages. Trolley problems are attractive laboratory tools because they provide easily modifiable stimuli to examine moral phenomena (Greene, Cushman, Stewart, Lowenberg, Nystrom, & Cohen, 2009). But, on the other hand, these dilemmas show a lack of ecological validity which is especially relevant for these types of studies, because moral cognition depends strongly on situational and cultural context (See Apendix III).

Box 1

Trolley Problem (Foot, 1967)

Suppose that a judge or magistrate is faced with rioters demanding that a culprit be found guilty for a certain crime and threatening otherwise to take their own bloody revenge on a particular section of the community. The real culprit being unknown, the judge sees himself as able to prevent the bloodshed only by framing some innocent person and having him executed. Beside this example is placed another in which a pilot whose airplane is about to crash is deciding whether to steer from a more to a less inhabited area. To make the parallel as close as possible it may rather be supposed that he is the driver of a runaway tram which he can only steer from one narrow track on to another; five men are working on one track and one man on the other; anyone on the track he enters is bound to be killed. In the case of the riots the mob has five hostages, so that in both examples the exchange is supposed to be one man's life for the lives of five.

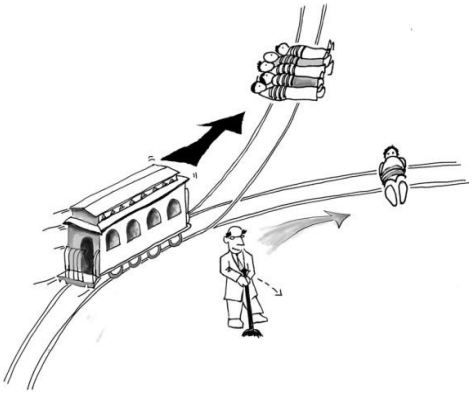


Figure 1. Footbridge Dilemma Illustration.

Box 2

Footbridge Problem (Tomson, 1985)

As before, a trolley is hurtling down a track towards five people. You are on a bridge under which it will pass, and you can stop it by putting something very heavy in front of it. As it happens, there is a very fat man next to you – your only way to stop the trolley is to push him over the bridge and onto the track, killing him to save five. Should you proceed?

Figure 2. Trolley Dilemma Illustration.

But the main question remains unclear. Why most people are willing to sacrifice one person to save five in the Trolley problem but not in the Footbridge dilemma? (Greene, et al., 2001; Hauser, Cushman, Young, Kang-Xing Jin, & Mikhail, 2007; Mikhail, 2007).

An answer came from the DPMs that contrast rational and emotional processing (Greene, et al., 2001; Greene, Nystrom, Engell, Darley, & Cohen, 2004; Greene, 2007; Shentav, & Greene, 2010; Cushman, Young, & Greene, 2010; Paxton, & Greene, 2010; Cushman, & Greene, 2012; Shentav, & Greene, 2014). These models predict that harmful actions in dilemmas like Footbridge elicit a prominent negative emotional response that inclines people toward disapproval. In contrast, people tend to approve of the action in the Trolley case because, in the absence of a countervailing prominent emotional response, they default into a utilitarian mode of reasoning that favors trading one life for five (Figure 3).

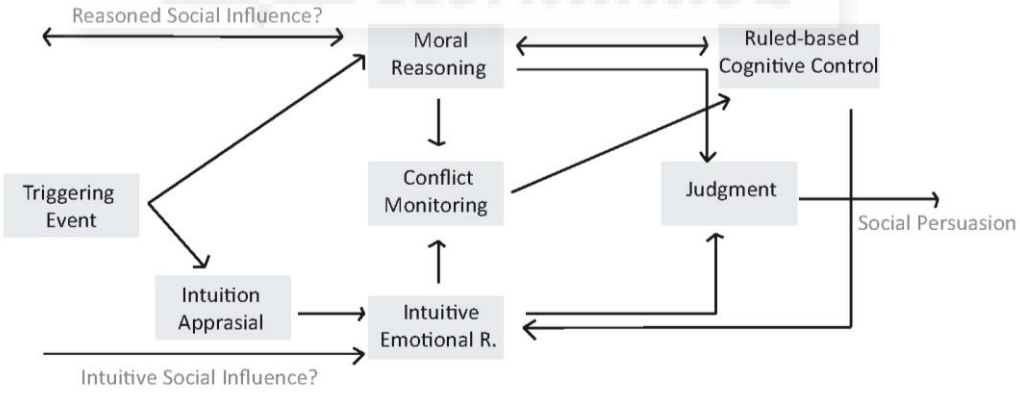


Figure 3. **The Dual Process Model Diagram.** According to Greene, moral judgments are driven by both intuitive emotional responses and controlled cognitive responses (from Paxton, & Greene, 2010).

Thus, DPMs explain why both dilemmas are judged differently, because the emotional system is more strongly engaged for the active "push" (Footbridge); and also why the push case seems to represent a more difficult dilemma than the switch case (Trolley) since a conflict between both systems is only triggered in the former (Greene, 2007; Greene, et al., 2009). In keeping with characteristic responses in the philosophical literature, the choice not to push is often referred to as "deontological", while the choice to push is denoted as "utilitarian".

Greene and colleges have also developed new moral test situations to dissociate and characterize the affective and rational processes that shape moral decisions (Greene, et al., 2001). In particular, three sets of dilemmas had been made, one with the relevant features of the Trolley dilemma, other with the relevant features of the Footbridge dilemma and the last without explicit moral content, i.e., neutral dilemmas. Dilemmas tailored from the Footbridge problem, in which the action would cause (a) serious body harm or death, (b) to a particular person or group, where (c) the harm does not result from deflecting an existing threat, were classified as personal, while the rest, with moral content, were classified as impersonal.

The personal/impersonal distinction is the most popular in the field of moral research (Cushman, & Greene, 2012). However, this classification has been also criticized for being overly crude and unable to explain the variability found in responses to the trolley problem (Mikhail, 2007). Greene, et al., (2004, 2009) have also been criticized for using more emotive language and references to family members or friends in their personal dilemmas and for failing to control for cognitive processing requirements across conditions (Borg, Lieberman, & Kiehl, 2008).

Thus, other classifications have been introduced depending on different factors:

- The distinction between active and passive harm (Cushman, Young, & Hauser, 2006; Baron, & Ritov, 2009; DeScioli, Bruening, & Kurzban, 2011). People generally consider it morally worse to harm a person actively than to passively allow a person to suffer harm.

- Harm as means versus harm as side effect (Mikhail, 2000; Royzman, & Baron, 2002; Hauser, et al., 2007; Cushman, & Young, 2010). Philosophers refer to this distinction as "doctrine of double effect" (Foot, 1967; Thomson, 1985). This dimension can be isolated from motor properties of pushing a person versus flipping a switch, and it continues to exert a substantial influence on moral research (Cushman, Young & Houser, 2006; Hauser, et al., 2007; Greene, et al., 2009).

- And the variables proposed by Christensen & Gomila, 2012: intentionality (whether the harm is done intentionally or instrumentally); evitability (whether the harm was avoidable or not); and benefit recipient (whether the profit had been obtained directly from the harm or not).

In any case, several sources of evidence support DPM hypothesis. Some studies undermine the role of controlled cognition: fMRI reveals correlates of controlled cognition for utilitarian choices (Greene, et al., 2004; Cushman, Murray, Gordon-McKeon, Wharton, & Greene, 2012), time pressure and cognitive load decrease the frequency and speed of utilitarian choices (Suter, & Hertwig, 2011; Trémolière, De Neys, & Bonnefon, 2012). Others underscore the role of affect: brain damage to regions that process emotions increases utilitarian responses, (Ciaramelli, Muccioli, Ladavas, & di Pellegrino, 2007; Koenigs, Young, Adolphs, Tranel, Cushman, Hauser, & Damasio, 2007; Moretto, Làdavas, Mattioli, & di Pellegrino, 2009) and people who exhibit low levels of affective concern for others make more utilitarian judgments (Bartels, & Pizarro, 2011; Gleichgerrcht, & Young, 2013).

Additionally, people with higher working memory capacity and those who are more deliberative thinkers are more likely to judge a harmful utilitarian action as permissible (Bartels, 2008; Feltz, & Cokely, 2009; Moore, Clark, & Kane, 2008). These findings suggest the influence of competing processes in personal dilemmas.

A key piece of behavioral evidence from that provides support for the DPM was the Reaction Time (RT) interaction between type of response (deontological vs. utilitarian) and the type of dilemma (personal or impersonal) (Greene, et al. 2001). Specifically, responses affirming the moral appropriateness in personal dilemmas were associated with longer RTs, but there were no differences for impersonal dilemmas. This difference in RTs, together with neuroimaging data, has been interpreted as an evidence that distinct neural sub-systems underlie utilitarian and deontological moral judgments.

This difference in RTs for personal dilemmas could not always be conclusively explained. McGuire, Langdon, Coltheart, & Mackenzie (2009) pointed out a flaw with respect to how the item analysis was performed in Greene, et al., (2001). In fact, repeating the original RT analysis they obtained a very different result. They did not find statistical differences in terms of RTs between responses and type of dilemma, concluding that, in this regard, there is not yet sufficient evidence to support the distinction between personal and impersonal moral dilemmas, and hence these distinctions cannot be used to support a DPM of moral judgment either (McGuire, et al., 2009).

Taken together, last previous reports seem to suggest that the simple division between “rational” and “emotional” processes is incomplete (Moll, Oliveira-Souza, & Zahn, 2008; Cushman, Young, & Greene, 2010; Kvaran, & Sanfey, 2010). It had been common to think that emotion and reason are two forces pulling each other in opposite directions, but both types of processes may involve some affective content, in the sense that they do not merely process information but yield competing motivations towards distinct behaviors. For instance, a system which supports utilitarian responses should not only represent the fact that 5 lives is more than 1 life, it should also involve some type of affective information related to the fact that choosing to save 5 lives are BETTER than saving 1.

Thus, the simple division between emotion and reason might not suffice to explain the moral mind. This point is appreciated by both critics and supporters of DPM, arguing that affective and cognitive components must be integrated (Moll, Oliveira-Souza, & Zahn, 2008; Kvaran, & Sanfey, 2010; Crockett, 2013). What is then required for a comprehensive theory of moral decision making is a framework in which affective content modulates behavior (see section 4).

4. Affective Value in Moral Judgments: The Social Intuitionist Model

A revolution in the science of emotion has emerged in the last few decades, with the potential to create a paradigm shift in thinking about decision theories. The research reveals that emotions constitute the most powerful, persuasive and predictable drivers of decision making (Frijda, 1988; Phelps, 2006; Keltner, & Lerner, 2010; Phelps, Lempert, & Sokol-Hessner, 2014). Across different domains, important regularities appear showing that emotions influence judgments and choices, proposing an integrated approach of decision making that takes into account the role of emotional inputs (Loewestein, & Lerner, 2003; Keltner, & Lerner, 2010).

In recent years, the number of scholarly papers on emotion and decision making has increased rapidly (Lerner, Li, Valdesolo, & Kassam, 2015). This new approach tries to demonstrate that "emotion is not a unitary construct, but rather a compilation of component affective processes that influence choices" (Phelps, Lempert, & Sokol-Hessner, 2014).

The term affect means the specific quality of "goodness" or "badness" experienced as a feeling state (with or without consciousness) and demarcating a positive or negative quality of a stimulus (Slovic, Finucane, Peters, & MacGregor, 2002) (for a more detailed description of the first evidences of affect see appendix IV). Affect is thus evoked by the simple presence of a perceptual experience, in this sense, it is one of the basic automatic parameter that impacts our daily life decisions (Zajonc, 1980; Slovic, et al., 2002; PhelpsLempert, & Sokol-Hessner, 2014). These types of judgments occur rapidly and become a guide for avoiding or, on the contrary, addressing decisions (Slovic, Finucane, Peters, & MacGregor, 2002; Slovic, et al., 2002).

The triggering event is labeled with different degree of affect in the process of decision making. Subsequently, people consult an "affect pool" which contains all the positive and negative tags consciously or unconsciously associated with the representations (Slovic, et al., 2002).

The features that become more salient in a judgment depend on characteristics of the individual (decision maker), the aspect evaluated (another person, object, circumstance, etc.), and the interaction between them. Individual differences in affective information processing systematically influence the information that matters most for the choice (Slovic, Finucane, Peters, & MacGregor, 2004). Affective processes appear to serve a default, background regulatory function responsible for guiding choices. Jonathan Haidt would later refer to this as "the emotional tail that wags the rational dog" (Haidt, 2001).

How affective reactions impact decisions is currently being widely investigated. This is particularly so in the moral field. For example, Haidt postulated that people often arrives at moral judgments through automatic processes and then use the controlled processes to construct an explicit rationale (Haidt, 2001). The key finding is that many people consider certain situation wrong but few can say precisely why (Haidt, Bjorklund, & Murphy, 2000; Haidt, & Hersh, 2001; Haidt, 2001). A potential explanation for this mismatch is that often

judgments depend on intuitions that directly value actions, while justifications are produced by post-hoc mechanisms of controlled reasoning (Haidt, 2001).

Haidt's Social Intuitionist Model (SIM) is composed of four principal links or processes. Every link is established by prior research in some domains of judgment. The model was presented as a proposal, up to now, no experimental data could unambiguously demonstrated the existence of these links (figure 4).

- Link 1. The intuitive judgment link. The model proposes that moral judgments appear in consciousness automatically and effortlessly as the results of moral intuitions (Zajonc, 1980; Bargh, & Chartrand, 1999; Greenwald, & Banaji, 1995).

- Link 2. The post hoc reasoning link. The model proposes that moral reasoning is an effortful process, engaged in after a moral judgment is made, because a person is searching arguments that will support the judgment (Nisbett, & Wilson, 1977; Kunda, 1990).

- Link 3. The reasoned persuasion link. Moral reasoning is produced and sent forth verbally to justify one's moral judgment. This reasoning can sometimes affect other people (Edwards, & von Hippel, 1995; Shavitt, 1990).

- Link 4. The social persuasion link. The mere fact that people have made a moral judgment exerts a direct influence on others, even if no outward conformity is issued (Asch, 1956).

These four links form the core of the SIM. The full model includes two ways in which private reasoning can shape moral judgments.

- Link 5. The reasoned judgment link. People may at times reason their way to a judgment by sheer force of logic, overriding their initial intuition (Wilson, Lindsey, & Schhooler, 2000).

- Link 6. The private reflection link. In the course of thinking about a situation a person may spontaneously activate a new intuition that contradicts the initial intuitive judgment (Selman, 1971).

Haidt began to extend the Social Intuitionist Model to specify the most important categories of moral intuition. The result was Moral Foundations Theory which was created to understand why morality varies so much across different cultures (Haidt, & Joseph, 2004; Graham, Nosek, & Haidt, 2011). The theory posits that there are (at least) five innate moral foundations, upon which cultures develop their various moralities, just as there are five innate taste receptors on the tongue (Haidt, & Joseph, 2004), which cultures constructs virtues, narratives, and institutions on top of these foundations, thereby creating the unique moralities that could be seen around the world. Table 1 shows the five modules and the emotions and virtues associated with them (Haidt, & Joseph, 2004).

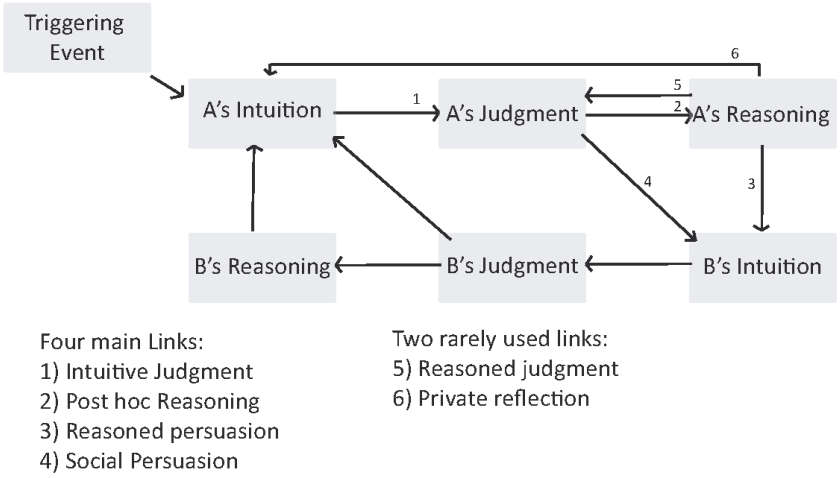


Figure 4. **The Social Intuitionist Model Diagram.** According to Haidt's moral judgments are driven mainly by Intuitions. Reasoning is usually produced after a judgment is made, in order to generate an explanation or influence other people. (From Haidt, 2001).

For Haidt, one possibility is that moral intuitions are the output of a small set of these cores. When a module takes the conduct or character of another person as its input and then emits a feeling of approval or disapproval, that output is a moral intuition.

Even though the SIM offers a good theoretical portrait of human morality, a number of critics have pointed out that intuitive judgments may reflect the 'automatization' of judgments based on prior moral reasoning (Saltzstein, & Kasachkoff, 2004), and that moral reasoning can disrupt the automatic process of judgment formation described by the SIM, either by a slow, intentional, deliberative and effortful 'after-the-fact' correction, or by an 'up-front' preconscious control (Fine, 2006). Thus new experimental evidence is needed to verify how exactly intuitions are implemented to make moral decisions.

Table (1)
The Five Moral Domains and the Emotions and Virtues associated with them (from Haidt, & Joseph, 2004)

	Harm/Care	**Fairness/ Reciprocity**	**Intergroup/ Loyalty**	**Authority/ Respect**	**Purity/ Sanctity**
Characteristic emotions	Compassion	Anger, gratitude, guilt	Group pride, belongingness, rage at traitors	Respect, far	Disgust
Relevant virtues	Caring, kindness	Fairness, justice, honesty, trustworthiness	Loyalty, patriotism, self-sacrifice	Obedience, deference	Temperance, chastity, piety, cleanliness

5. DPM & SIM

Currently most studies of moral decision making propose that moral judgments depend on two kinds of cognition: moral intuition and moral reasoning. Moral intuition refers to fast automatic and affect-laden processes in which an evaluative feeling of good-bad or like-dislike (about the actions or character of a person) appears in consciousness without any awareness of having gone through steps of search, weighing evidence, or inferring a conclusion (Haidt, 2001; Greene, et al., 2001). Moral reasoning, in contrast, is a controlled process; it is conscious mental activity that consists of transforming information about people and their actions in order to reach a moral judgment or decision (Greene, et al., 2001). It is suggested that depending on the type of dilemma the two processes would be differently engaged (Greene, et al., 2001).

The critical difference between Haidt's and Greene's proposals is that while the SIM posits that reasoned judgment occur in very few cases, just when the intuition is weak and processing capacity is high (Haidt, 2001), for Greene's DPM, moral reasoning, specifically, utilitarian reasoning, is a key aspect of our moral sense (Greene, et al., 2001, 2004; Cushman, Young, & Greene, 2010).

6. Canonical Mechanism

Finally, despite the progress that has been made in recent years and the agreement about certain aspects, like the role of affect, one of the key issues that need to be addressed is the algorithmic mechanism that underlies moral decisions. To this day, although extensive research has been carried out on the psychological bases of moral choices (Mikhail, 2007; Crockett, 2013), few studies have dealt with the nature of the computations that govern this decision process. Given that the exact underlying neural mechanism is still unknown, our aim is to propose which type of operation is performed when making a decision.

Many of our decisions, instead of being based on absolute valuations, rely on context dependent choices, where the election depends on the composition of the choice set (Tversky, & Simonson, 1993). Context should strongly influence how, for instance, the different values of luminance in the background, affect the perceived luminance of an object. This context-dependent modulation is precisely described by divisive normalization (DN), indicating that this standard form of sensory gain control may be canonical neuronal computation (Louie, Grattan, & Glimcher, 2011; Carandini, & Heeger, J., 2012; Louie, Khaw, & Glimcher, 2013; Louie, LoFaro, Webb, & Glimcher, 2014).

DN, firstly, has been proposed to be a general mechanism throughout the visual system. And not only that it is also the basic computation in other perceptual modalities, and cognitive systems in many species; for example it explains computations in the olfactory system of invertebrates (Olsen, Bhandawat, & Wilson, 2010), explains the light adaptation and surround suppression in areas ranging from the retina to cortex (Carandini, & Heeger, 2012), visual attention (Reynolds, & Heeger, 2009), multisensory integration (Louie, Grattan,

& Glimcher, 2011) and explains value-guided choice experiments in both monkeys and humans (Louie, Khaw, & Glimcher, 2013).

Normalization during the decision making process promotes a rescaling of neural activity (or the value associated with the stimuli) driven by the value of all choice options, implementing a relative value code (Louie, Khaw, & Glimcher, 2013).

A perfect example of context dependent choices could be moral decisions. This type of choice behavior, frequently, violates the fundamental norms of rational theories suggesting that the exact nature of the moral response depends on the relative valuations assigned to individual options.

Therefore, we hypothesize here that moral decision making is shaped by the same gain control mechanisms that operate in sensory pathways to represent value in a relative rather than absolute manner. Context should thus strongly influence how people judge actions and actors in a very explicit way that is independent of purely deontological or absolute utilitarian considerations.

Aims

The main motivation of this PhD thesis was to propose a mechanism that could explain the neuronal operations underlying moral decisions. To that end we have used computational techniques informed by a large set of behavioral recordings in four different experiments. In particular, we set out to address the following questions:

1. Mechanism in Moral Decision Making:
   - 1.1 How the human brain makes moral decisions? What mechanism underlies morality?
   - 1.2 How do individual differences contribute to the moral decision making process?
2. Morality at the Population Level:
   - 2.1 Do we all have a similar way of approaching a moral dilemma?
   - 2.2 Have men and women different moral behaviors?
   - 2.3 Have men and women different affective perceptions?
3. Dynamics of Moral Decisions:
   - 3.1 Have the different behaviors a specific cognitive load[1] pattern?
   - 3.2 Have the different behaviors a distinct timing?
4. Can moral behaviors be predicted?
5. Awareness of Moral Decisions:
   - 5.1 Are moral decisions unconscious?
   - 5.2 Have the different moral behaviors a specific degree of awareness?
6. Do we use different strategies to solve dilemmas with different degrees of difficulty?

[1] Cognitive load is defined as the effort associated with the task, as introduced by John Sweler in the late 1980s (Sweler, 1988).

Materials & Methods

This chapter describes the materials and methods used for the data collection and posterior analysis. The section is divided into four different parts. The first subsection explains in detail the different experiments that were conducted. The aims of these experiments were to evaluate the following variables:

- Stereotyped Affect (SA): a basic reaction about the likeability of the active character on the dilemma.
- Perceived Affect (PA): this variable is the same as the previous one, but in this case, participants judge the protagonist, not in isolation, but when she/he was involved in a given situational context.
- With both measures, we obtained indirectly the value of the Contextual Affect (CA) (For a more detailed definition see Appendix V).
- Acceptability (Acc): whether, the action depicted in the dilemma was deemed morally acceptable or not.

The second part moves on to describe in greater detail how the data analysis was performed. And, the third and fourth sections specify how the behavioral and statistical analyses were performed, respectively.

Appendix VI shows a particular psychometrical analysis.

Finally, appendix VII and annexes 1 through 7 provide more information about the materials and questionnaires that were used in these experiments.

1. General Experimental Procedure

First of all participants gave their informed consent (Annex 1) before taking part in the experiments. Instructions were given both orally and in writing. We also requested that subjects provided demographic data such as age, gender, qualifications, religion affiliation, visual problems and if they were right or left handed.

All the experiments were approved by the University's Ethical Committee of the Miguel Hernández University (code: IN.LMO.02.15). Moreover participants who performed the experiments were naïve to the research project. We ensured the confidentiality and anonymity of the responses and scientific purposes of the study.

1.1 Experiment 1

1.1.1 Participants

We recruited 14 participants (10 female, 4 male; mean age 28.38 ± 4.50 SD), the study was run with graduate students of the Institute of Neuroscience, a joint institution of the Spanish National Research Council and the Miguel Hernández University.

1.1.2 Stimuli

A set of 19 personal stereotypes were presented (Annex 2). The first five items were training questions. Some of these characters will then appear as the protagonists of the moral dilemmas used in the rest of the experiments.

1.1.3 Procedure

Specifically, participants were asked to evaluate the personal SA, responding to this question "How much do you like this character?" Using a 5-point likert scale (1= low SA, 5= high SA).

1.2 Experiment 2

1.2.1 Participants

We recruited 16 participants (10 female 6 male; mean age 26 ± 3.77 SD). (The study was run with students of the Institute of Neuroscience in the Miguel Hernández University of Elche, as in the previous section). These subjects were not involved in the previous study.

1.2.2 Stimuli

A total of 10 moral dilemmas were used (Annex 3). The moral items were real stories, which have appeared in spanish newspapers over the past few years, with the exception of one of them which was adapted from the popular quandary used by Kohlberg during his experiments, the Heinz dilemma (Kohlberg, 1958). The reason for using this new material was to increase the ecological validity of the moral situations that are usually used in this field. Each moral dilemma was divided into four trials. The first trial always exposed the main facts with as little contextual information as possible, which was then sequentially introduced in the following ones, with the intention to modify the emotional content of the moral scene (Appendix VI shows a special psychometrical analysis to evidence the inter-relatedness of the items within the experiment). Moreover, half of the dilemmas were written in the first person, and half of them included an uncertain, probabilistic outcome.

1.2.3 Procedure

Participants were asked to evaluate the PA of the protagonist who appeared in every single trial that composed the whole moral scene.

Subjects were instructed to respond, as fast as possible after hearing the trial the following question: “How much do you like this character in the given situation?” using a 1-5 point likert scale (1= low PA, 5= high PA) (Annex 4).

1.3 Experiment 3

1.3.1 Participants

A total of 37 subjects (18 female, 19 male; mean age 25.65 ± 4.57 SD) took part in this experiment. Participants were recruited through the student-based and public subject pools at the University of Leicester. They were compensated for their time and travel with 10£. Data from two participants were excluded due to calibration problems during data collection. All of them reported normal or corrected-to-normal vision. All participants were fluent speakers of either Spanish or English. The experiment could be run both, in Spanish (45.7%) and English (54.3%).

First, all participants completed a psychological questionnaire in order to exclude those subjects who don’t have the proper mental state for taking part on a moral experiment. We used the LSB-50 (de Rivera, & Abuín, 2012) for Spanish speaking subjects and the BSI-56 (Derogatis, 1992) for English speaking participants. (For a more detailed description of the questionnaires and the results, please consult Appendix VII).

1.3.2 Stimuli

We used the same contextual moral dilemmas previously used in experiment 2. As a control, we used two kinds of questions: yes/no simple questions and visual illusions (Annex 5). The first type of control questions were composed of a total of 9 items. These trials were divided into 3 groups depending on the type of answer. The first group included questions for which the correct response was “Yes”. The second one, the correct response was “No”. And, finally, during the three remaining trials, participants were informed specifically that they had to lie.

Each visual illusion was composed of three trials. The first and the third ones were the same, while on the second one, the visual context was different. Participants were asked about perceptual differences between visual attributes, such as, color, shade of grey, etc. in the different contexts.

1.3.3 Recordings

To evaluate the dynamic aspects of a moral decision we continuously monitored the gaze of the participants and their hand’s motion while responding to our moral dilemmas. Please see Annex 6 for a real example.

1.3.3.1 Eye Tracking Recordings

Eye tracking was performed using an EYELINK 1000 eye tracker (Figure 5A) running at 500Hz on a CRT monitor with a resolution of 1024 x 768, 32 bit color depth, and a refresh rate of 70 Hz. Stimuli were presented using PsychoPhysics toolbox running on Matlab 2013a, with a viewing distance of 57 cm. Eye position was tracked at 500 Hz with an average noise (RMS value) <0.010 (Figure 5B). All eye movements were labeled as fixations, saccades, and blinks by the eye-tracker software using the default thresholds for Psychophysics experiments (30 deg/sec for velocity, 9500 deg/sec^2 for acceleration, and 0.15 deg for motion; Cornelissen, Peters, & Palmer, 2002).

Before starting an experiment, the eye-tracker was calibrated with a set of 5 points. For the observer, the calibration stage simply consisted of fixating the gaze at the 5 points displayed sequentially and randomly at different locations on the screen (Figure 5C). The validation of the calibration was achieved by measuring the difference between the computed fixation positions and the locations of the target points. Such a difference reflects the accuracy of the eye movement recording. A threshold error of 1^0 has been selected as the greatest divergence that could be accepted. The participants were seated on an adjustable height chair so that the eyes of all subjects could be aligned with the center of the screen. We have used a head holder that allows flexibility to perform the experiment in free viewing conditions (Figure 5D), so the recorded data is not affected by head motion and allows observers to view stimuli normally with no restrictions imposed by head movements. The experiment was performed in a dark room.

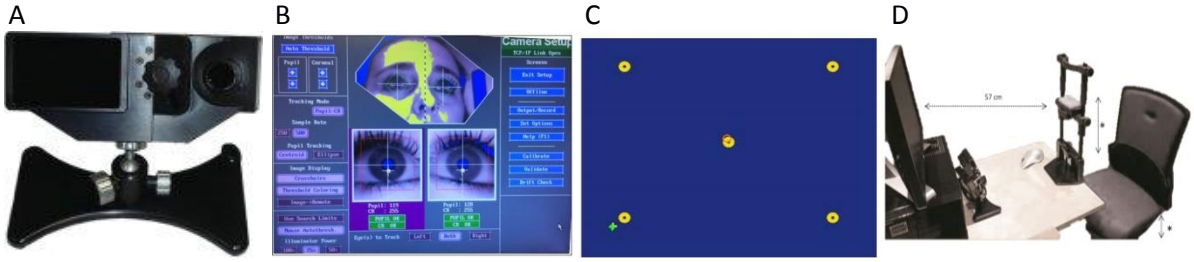


Figure 5. Set up. A) EyeLink used for recording the fixation and saccadic eyes' movements**.** B) EyeLink settings display. C) Calibration display. D) Environmental conditions.

1.3.3.2 Mouse Tracking Recordings

The mouse tracker was developed using Psychotoolbox. Data were recorded at 1000 Hz. While the audio containing the questions was on, the mouse was locked and positioned inside the Start Box (Figures 6, 7 & 8).

1.3.4 Procedure

There were a total of 28 dichotomous quandaries (10 moral dilemmas, 9 visual illusion related questions and 9 simple, control questions) that participants were asked to listen to and make judgments about. The experiment was designed carefully so that it should not exceed 45 minutes including the calibration stage, in order to prevent the observers' loss of interest or disengagement from the task over time.

Participants were told about the study, detailing what the experimental procedure would consist of. At the beginning of the experiment they would hear 9 simple questions (Figure 6). These initial trials, worked both as control question and as training since the location of the YES - NO boxes remained unchanged throughout the experiment. The left one would be the YES box (coordinates 51.50 - 77; 256.5 - 192) and the right one the NO box (coordinates 767.5 – 77; 972.5 - 192).

After these trials, subjects had to answer questions about moral dilemmas and visual illusions interleaved. In the case of moral dilemmas, the question was always the same: Is this behavior morally acceptable in this context? This question appeared in the center of the screen, so the participants could read it before hearing the moral dilemma. When they were ready, they pressed the start button (coordinates 437 - 653; 587 - 730) with the mouse to start the audio track. While the audio was playing, the participants fixed their gaze in the black cross, which appeared in the center of the screen (coordinates 512 - 284) (Figure 7). Participants were instructed to respond as quickly as possible after hearing the dilemma.

All the dilemmas were played only once. The alternatives would be visible after the audio without limited time. The YES and NO boxes had the same position than during control questions. After every other trial, a 1-5 continuous scale was presented to evaluate the participant's confidence on her own response. For more detailed information see figure 7.

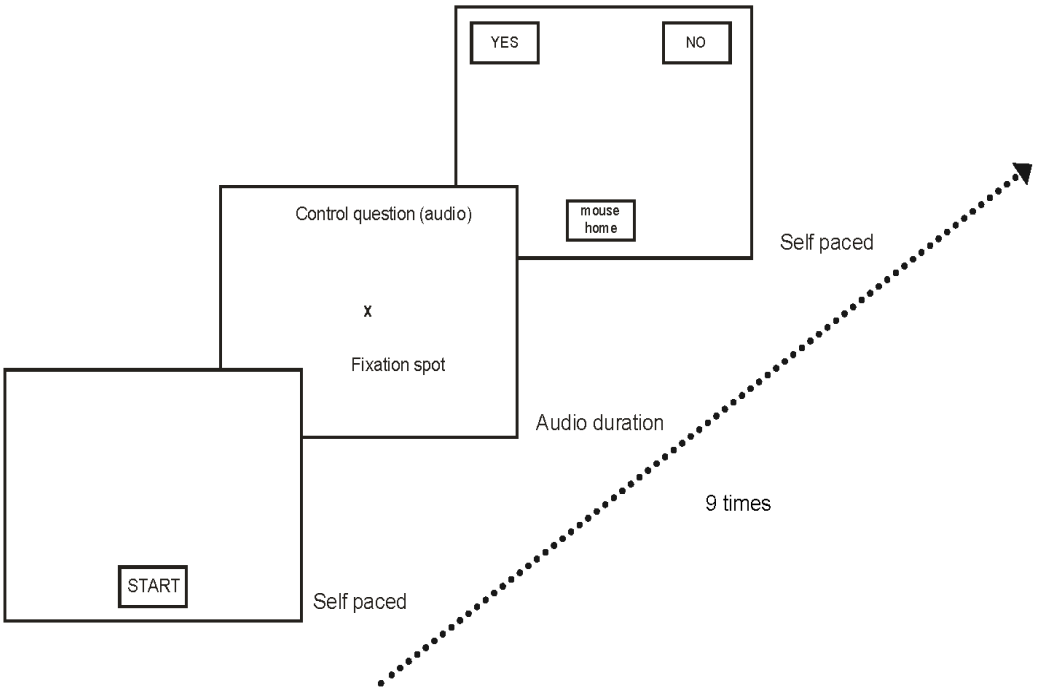


Figure 6. Experimental Design, Control Questions. In the first slide subjects decide when they are ready to start the experiment. During the second screen the participants hear the control question while the eyes are fixed in the central cross. The third screen prompts the subject to respond moving the mouse and clicking in one of the two boxes. Subjects had unlimited time to respond, although they were instructed to do so as quickly as possible.

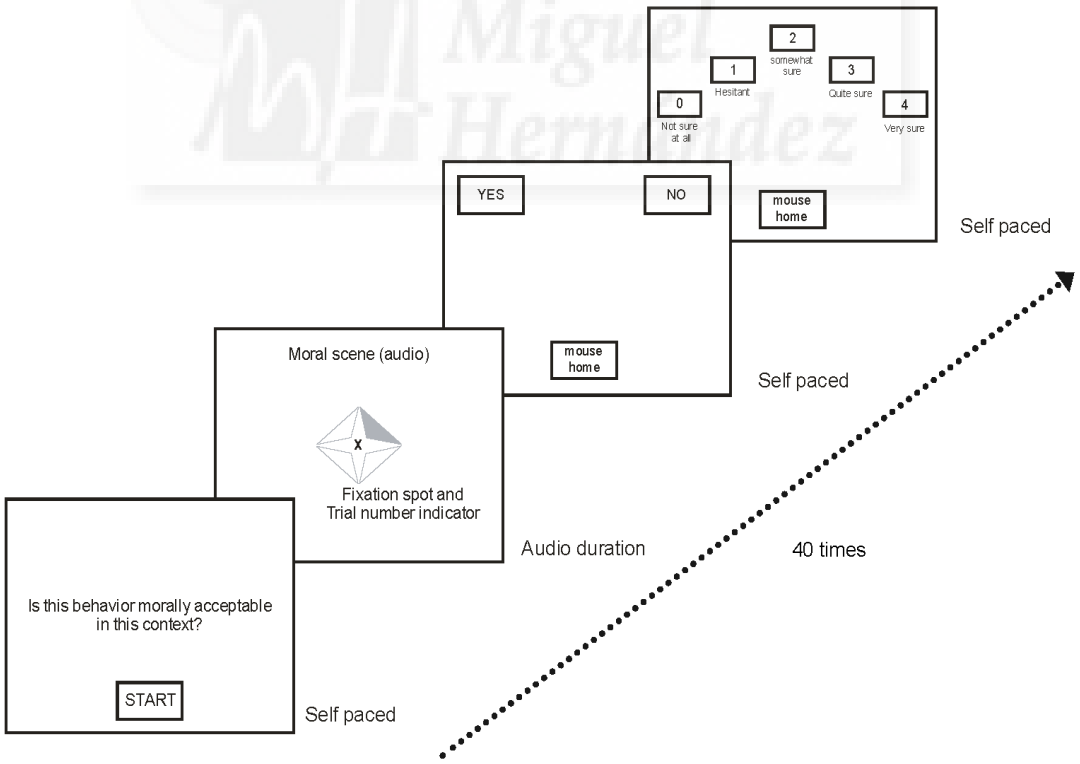


Figure 7. **Experimental Design, Moral Task.** In the first slide subjects decide when they are ready to start the experiment after reading the question they must respond. During the second screen the participant hears the moral dilemma while her eyes are fixed in the central cross. A figure appears in the center of the screen to indicate the trial's number 1-2-3-4. The third screen prompts the subject to respond by moving the mouse and clicking in one of the two boxes. Subjects had unlimited time for responding although they were instructed to do so as quickly as possible. The fourth screen with the confidence scale only appears in every other dilemma.

Each visual illusion was composed of three audio tracks and appeared after the performance of two moral dilemmas. Participants had to hit the start button to begin each run. After that, a visual illusion would appear in the center of the screen while the subjects heard the question. As in the case of the moral scenes, the response alternatives would only be visible after the audio was finished. In some cases, the boxes contained de usual YES and NO answers, and in other cases EQUAL (right) and DIFFERENT (left). After every other trial, a 1-5 continuous confidence scale was presented (Figure 8).

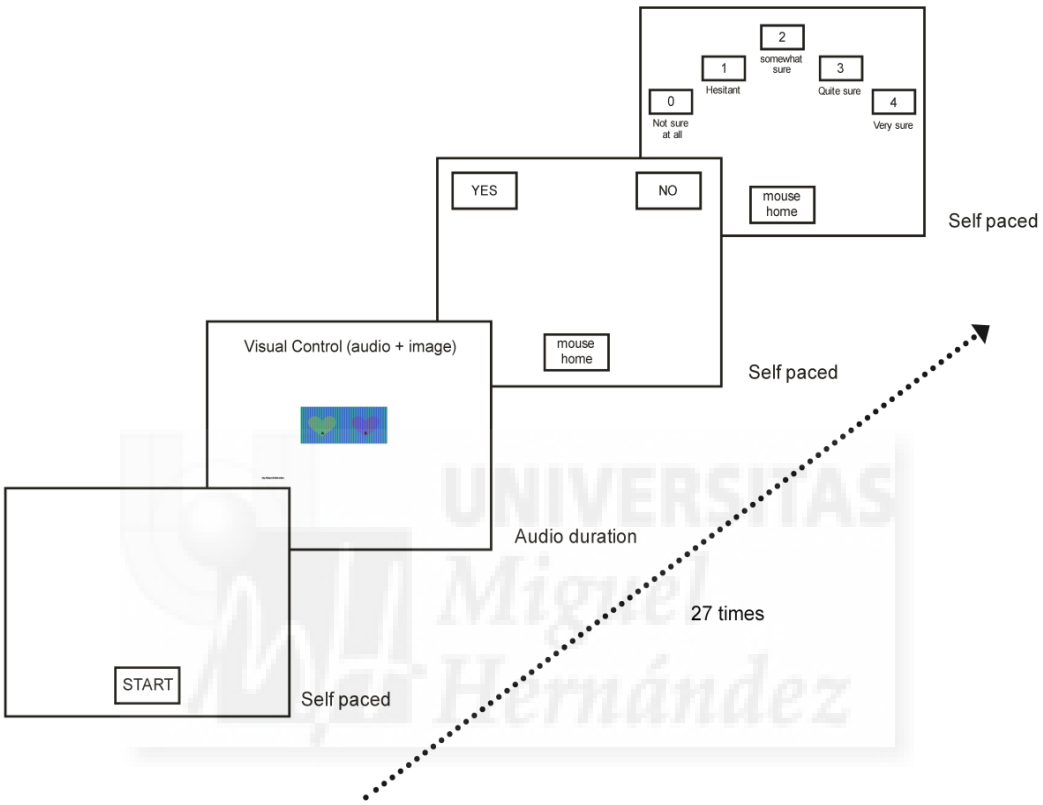


Figure 8. Experimental Design, Visual-Control Task. In the first slide subjects decide when they are ready to start the experiment. During the second screen, the participant sees the illusion, which appears in the central of the screen, while she is hearing the question. The third screen prompts the subject to respond by moving the mouse and clicking in one of the two boxes. Subjects had unlimited time for responding although they were instructed to do so as quickly as possible. The fourth screen with the confidence scale only appears in every other illusion.

The moral scenes were randomized and divided in different blocks with the same inner structure. On each block, participants performed nine control trials followed by two moral scenes and one visual illusion.

1.4 Experiment 4

1.4.1 Participants

We recruited 442 participants (314 female, 128 male; mean age 21.86 ± 4.456 SD). The study was run with undergraduate students of the Miguel Hernández University of Elche.

1.4.2 Stimuli

We used the same moral dilemmas we had used in experiments 2 and 3.

1.4.3 Procedure

Participants were asked to listen to each of the trials that comprised each moral dilemma and to answer the following questions.

- The PA of the protagonist of each moral scene, using a 5-point Likert scale (1= low PA, 5= high PA), in response to the question "How much do you like the main character in the dilemma?"
- Whether or not the action was morally acceptable in each context (Acc).
- The level of confidence with which they answered the questions, using a 5-point Likert scale (1= low confidence, 5= high confidence).

We used six different types of questionnaires (Annex 7). Each questionnaire contained three moral scenes. Each participant responded randomly to two or three scenes.

2. Data Analysis

All variables (SA, PA and Confidence Report (CoR)) were recorded using discrete, 1-5, Likert scales. SA and CoR were later transformed into continuous 0-1 scale, where 1 will be 0 and 5 will be 1; while PA was later transformed into -1 to 1 scale, where 1 will be -1 and 5 will be 1. The motivation of this transformation is due to PA is defined as the contrast between SA and a variable that will be introduced below, namely, the Contextual Affect or CA. This will be explained in the next section.

2.1 Divisive Normalization

The normalization model (Carandini, & Heeger, 2012) is here defined by an equation which specifies how moral decisions, i.e. the probability of accepting as moral a given action in a particular situation, depend on choice context. Normalization performs a rescaling of the input driving by the value of all choice options, implementing a relative value code (Louie, Grattan, & Glimcher, 2011; Carandini, & Heeger, 2012; Louie, Khaw, & Glimcher, 2013; Louie, et al., 2014).

$$P(accep) = \gamma \frac{SA^n}{\sigma^m + (CA^m + SA^m)} - \beta$$

In our case, the numerator includes the main driving input, or Stereotyped Affect (SA). The denominator includes a constant *σ* plus the normalization factor, which is the sum of the Contextual Affect (CA) and SA, the normalization pool. The constants *γ*, *σ*, *β*, *m* and *n* constitute free parameters that are typically fit to empirical measurements: *γ* determines overall responsiveness; *σ* prevents division by zero and determines how responses saturate with increasing driving input, m and *n* are exponents that amplify the individual inputs and the constant *β* that sets the offset, simulating a baseline.

The CA is derived experimentally from PA and SA as follows:

$$CA = SA^{\frac{(1-PA)}{(1+PA)}}$$

In this way the PA is defined as the contrast between SA and CA as follows:

$$PA = \frac{(SA-CA)}{(SA+CA)}$$

2.2 Perceived Affect Contrast

Contrast values in our moral dilemmas were calculated using Michelson Contrast (Takeuchi, & De Valois, 2000) as follows:

$$PA\ Contrast = \frac{(MaxPA - MinPA)}{(MaxPA + MinPA)}$$

where MaxPA and MinPA stand for maximum and minimum PA, respectively. To compute MaxPA and MinPA it was necessary to, first, assign the response from each subject to each of the dilemmas to a different behavioral group (D, C1, C2, C3 or C4; see behavioral analysis below), depending on whether the subjects had accepted the action as moral or not and, if so, the number of trial in which the acceptance had occurred. Second, we computed an average PA per trial and used it to obtain a PA contrast for each cluster in each dilemma. PA Contrast was, therefore, the average of the individual PA contrast scores.

2.3 Acceptability Index

Acceptability Index (AI) was calculated for each subject (only for participants of experiment 3 since they were the only ones who respond to all the dilemmas) from the participants' response in every single trial. Trials with Contextual Response (CR; accepted response) were coded as 1 and trials with Deontological Response (DR) were coded as 0. An average intra subject was performed.

$$AI = \frac{\sum CR}{n}$$

Where n was the total number of responses per participant.

2.4 Doubt Index

Doubt Index level (DI) was determined using the eye tracker recordings and different measures of time that participants consumed in each moral decision.

$$DI = 1 - \frac{Ti}{RT - \left(\frac{To}{RT}\right)1s}$$

Where Ti is the general average time, in seconds, of the 6 positive and negative control questions (M = 1.6030; SD = 1.2554), RT is the total time consumed in each trial and To is the time that a participant had spent out of the response box in the specific trial. It is bound between 0 and 1, where values closed to 0 indicate low doubt, while values closed to 1 high doubt.

2.5 Linear Discriminant Analysis

Linear discriminant analysis is a method to find a linear combination of features that characterizes or separates two or more classes of objects or events and, in our case, both the response (cluster where the participants belong to) and the dilemma that participants were responding to. Decoding performances are usually quantified by the relative number of hits that are the average of the diagonal in the confusion matrix (Quian Quiroga, & Panzeri, 2009; Navajas, Ahmadi, & Quian Quiroga, 2013). As the outcomes of the predictions of each stimulus

can be regarded as a sequence of Bernoulli trials (independent trials with two possible outcomes: success and failure), the probability of successes in a sequence of trials follows the Binomial distribution. Given a probability *p* of getting a hit by chance (*p = 1/K*, in which *K* is the number of stimuli), the probability of getting *k* hits by chance in *n* trials is given by

$$P(K) = \binom{n}{k} p^k (1-p)^{n-k}$$

Where

$$\binom{n}{k} = \frac{n!}{(n-k)!\,k!}$$

is the number of possible ways of having *k* hits in *n* trials. From this it is possible to assess statistical significance and calculate a *p*-value by adding up the probabilities of getting *k* or more hits by chance:

$$\text{pvalue} = \sum_{j=k}^{n} P(j)$$

First of all, trial 2, 3 and 4 were normalized at the first trial. To cross-validate decoding results, some trials were used as the training set. This procedure was the "leave-one-out" in which each trial is predicted based on the distribution of all the others trials.

2.6 Confidence Report

Confidence Reports (CoR) were obtained directly from the scores provided by participants using the Likert scale after answering the moral trials.

2.7 D-prime

Using the paradigm of the signal detection theory, different self-awareness responses were mapped (Quian Quiroga, & Panzeri, 2009). The DI values were employed as signal and the CoR were used as classifier. The DI and CoR were calculated for the participants of experiment 3. First of all it was necessary to calculate the CoR median for each participant. The median was used to determine the threshold value from which CoR were classified as low, if the scores were smaller than the median, or high if the values were larger.

The four possible outcomes were formulated in a 2×2 contingency table (Table 2), in which they were labeled either as positive or negative, as follows:

If the outcome from a prediction was positive and the classifier value was also positive, then it was called a true positive (TP); however if the classifier value was negative then it was said to be a false positive (FP). Conversely, a true negative (TN) occurred when both, the prediction outcome and the classifier value were negative, and false negative (FN) was when the prediction outcome was negative while the classifier value was positive. Table (3) summarizes how the rates were calculated.

Table (2)
Signal Detection Theory: Confusion Matrix

	Condition Positive (High DI)	**Condition Negative (Low DI)**
Test outcome positive (Low CoR)	True Positive (TP)	False Positive (FP)
Test outcome negative (High CoR)	False Negative (FN)	True Negative (TN)

The sensitivity index or *d'* was also calculated. It provides the separation between the means of the signal and the noise distributions, compared against the standard deviation of the noise distribution. d' is defined as:

$$d' = Z(\text{True Positive Rate}) - Z(\text{False Positive Rate})$$

Where function $Z\ (p)$, $p \in [0, +1]$, is the inverse of the cumulative distribution function of the Gaussian distribution.

Table (3)
Signal Detection Theory: Rates

Index	Formula	Index	Formula
Negative Predictive Value	$\frac{\sum TN}{\sum Test\ outcome\ Negative}$	Prevalence	$\frac{\sum Condition\ Positive}{\sum Total\ Population}$
True Positive Rate	$\frac{\sum TP}{\sum Condition\ Positive}$	Positive Predictive Rate	$\frac{\sum TP}{\sum Test\ outcome\ Positive}$
False Negative Rate	$\frac{\sum FN}{\sum Condition\ Positive}$	False Omission Rate	$\frac{\sum FN}{\sum Test\ outcome\ Negative}$
True Negative Rate	$\frac{\sum TN}{\sum Condition\ Negative}$	Accuracy	$\frac{\sum TP + \sum TN}{\sum Total\ Population}$
False Positive Rate	$\frac{\sum FP}{\sum Condition\ Negative}$	False Discovery Rate	$\frac{\sum FP}{\sum Test\ outcome\ Positive}$

2.8 Rasch Model

The Rasch model is a psychometric model for analyzing categorical data, as a function of the trade-off between the respondent's abilities and the item difficulty (Baron, Gürçay, Moore & Starcke, 2012).

In the Rasch model, the probability of a specific response (deontological or contextual (considered wrong in this case)) is modeled as a function of person and item parameters, the probability of a correct response is modeled as a logistic function of the difference between the person and item parameters.

Let x_{ni}= x ∈ [0, 1] be a dichotomous random variable where x = 1 represents a correct response and x = 0 an incorrect response to a given assessment item. In the Rasch model for dichotomous data, the probability of the outcome x_{ni}= 1 is given by:

$$\mathrm{P(a)} = \frac{e^{(\theta-\beta)}}{1+e^{(\theta-\beta)}}$$

Where β is the ability of person and θ is the difficulty of item, both in a logit scale.

$$\text{Logit (x)} = \ln\left(\frac{x}{1-x}\right)$$

A Difficult Index (DifI) was obtained for each dilemma.

$$DifI = 1 - \frac{\sum CR}{n}$$

An Ability Index was computed for each participant. This index was equal to the Acceptability Index (AI).

$$AI = \frac{\sum CR}{n}$$

Statistic Infit was calculated to quantify the model's fitting.

$$\mathrm{Infit} = \frac{\sum \mathrm{Z_{is}^2 W_{is}}}{\sum \mathrm{W_{is}}}$$

Where Z is the standardized residuals, W is the variance, i refers to the specific trial and s denotes the subject.

Z was obtained following the formula.

$$Z_{is} = \frac{Y_{is} - P(a)_{is}}{\sqrt{P(a)_{is}(1-P(a)_{is})}}$$

Where Y_{is} is the response of the subject to the dilemma.

Rasch fit statistics describe the fit of the items to the model. The mean square fit statistics have a chi-square distribution and an expected value of 1, where fit statistics greater than 1 can be interpreted as demonstrating more variation between the model and the observed scores. Values of Infit less than -2 or greater than +2 usually indicates that the item or person have a poorer compatibility with the model.

3. Behavioral Analysis

3.1 Behavioral Clustering

We assigned the responses of each subject to each of the dilemmas to a different behavioral group (D, C1, C2, C3 or C4), or cluster, depending on whether the subjects deemed the actions depicted in the dilemmas as morally unacceptable, by clicking in the response box "NO" (responses labelled DR, as in deontological), or morally acceptable, by clicking in the response box "YES" (responses labelled CR, as in contextual); and, if the latter, the number of trial (1, 2, 3 or 4, hence C1, C2, C3 or C4) in which the acceptance first occurred (Figure 9). For simplicity, when a subject changed two or more times of response sign (acceptability) during the course of one dilemma, the data from that particular subject and dilemma was excluded from the final analysis (Experiment 3: 13 dilemmas of 322, 4%; Experiment 4: 83 dilemmas of 730, 11.3%).

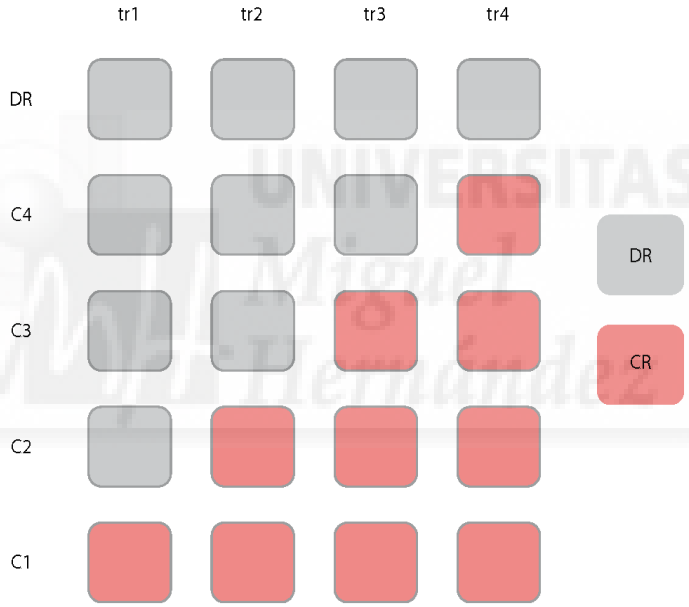


Figure 9. **Behavioral Clusters.** Behavioral classification of subjects depending on the type of response in the four trials of a dilemma.

3.2 Cognitive Load

To determine if the cognitive load associated to a moral decision predicts the different response patterns (or behavioral clusters described above) we analyzed the dynamics of the choice. First, in the course of experiment 3, we computed three parameters associated to the performance of each subject in response to each dilemma. a) RT measures the response time, or latency, elapsed between the end of the audio track and the response (mouse click on the response box). b) Number of visual fixations, measured using the eye tracker, made by the subject while exploring the different elements of the response screen (panel 3 in Figure 7) in the course of the response. And, c) the changes in acceleration experimented by the computer

mouse while the subject moved it from the start to the response box. We used these data to define a swift and a deliberate trial per dilemma and subject. Swift trials were those with the lowest scores for at least two of the three parameters (RT, visual fixation and mouse acceleration) considered above. Conversely, deliberate trials were those with the highest scores on at least two of those same three parameters. To minimize noise, the number of visual fixations performed during the first 300 ms of the response (directed mostly to the fixation cross in the center of the screen) and the mouse movements during the first and last 10% of the response time (dominated by minute displacements related to the launching of the movement and the landing on the target, respectively) were not included in the analysis.

4. Statistical Analysis

All the statistical analyses were performed using the software Prism 5 (GraphPad Software, Inc., California, USA), Matlab R2013a (The Mathworks Inc., Massachusetts, USA), STATA (Data Analysis and Statistical Software) and IBM SPSS Statistics 23 (Statistical Package for the Social Sciences).

4.1 Relation between variables

- Chi-square: We used this statistic (χ^2) to assess the strength of the relationship between the two categorical variables, the Acc and the PA. The aim was to determine whether the dependent variable (Acc) is contingent on the independent variable (PA) in every single sample.

- Spearman Correlation: We employed the Spearman coefficient (r_s) to measure the statistical dependence between the PA with the Acc; and the SA with Acc. The particularity of this analysis is that we have used the values of PA and SA calculating the average per trial; and the proportion of Acc per trial.

- Pearson Correlation: We employed the Pearson coefficient (r) to measure the statistical dependence between the recorded and predicted RTs'.

4.2 Reliability Analysis of the items

- Cronbach's Alpha (α): We calculated this index to provide a measure of the internal consistency of the trials; it is expressed as a number between 0 and 1. Internal consistency describes the extent to which all the items in a test measure the same concept or construct and hence it is connected to the inter-relatedness of the items within the test.

$$\alpha = \frac{N * \bar{c}}{\bar{v}(N-1)\bar{c}}$$

Here N is equal to the number of items, $\bar{c}$ is the average inter-item covariance among the items and $\bar{v}$ equals the average variance.

4.3 Proportion Analysis

- Chi-square: We used this statistic (χ^2) to test the equality of two proportions obtained from independent samples.

4.4 Statistical Modelling of non-Gaussian variables

- Generalized Linear Models (GLM_z): Our purpose of using this type of statistical modeling is to fit the observed data well, but also to be able to predict the future trend using the optimal independent variables. GLM_z are an extension of the linear modeling process but they extend the ideas of regression analysis to a wider class of problems involving the variables which do not follow the normal distribution, like in our case.

  GLM_z uses a link function which specifies a nonlinear transformation of the predicted values. Various link functions are commonly used, depending on the assumed distribution of the dependent variable values.

  Specifically we have used GLM_z to determine which independent variables were responsible for the RT (dependent variable). RT had a tweedie distribution, which are a family of probability distributions that include the purely continuous normal and gamma distributions, the purely discrete scale Poisson distribution, and the class of mixed compound Poisson-Gamma distributions which have positive mass at zero, but are otherwise continuous. Log link function is commonly used for this type of distribution. (McCullagh, & Nelder, 1982; McCullagh, & Nelder1989)

4.5 Non parametric test

- The Kruskal-Wallis (K) statistic is a nonparametric (distribution free) test, and it is used when the assumptions of ANOVA are not met. This statistic assessed for significant differences on a continuous dependent variable by a grouping independent variable (three or more groups).
- The Mann-Whitney test (U) is the alternative test to the independent sample t-test. It is a non-parametric test that is used to compare two samples that come from the same population, it is also used to test whether two subpopulation means are equal or not.

4.6 Post hoc comparisons

- For post-hoc analyses we performed statistical tests and corrected for multiple comparisons with Bonferroni's method.

4.7 Validation

- Cross-validation is a model assessment technique that we have used to evaluate the capacity of PA in making predictions on new datasets that it has not been trained on. This is done by partitioning a dataset (k=5) and using a subset to train the algorithm and the remaining data for testing. Because cross-validation does not use all of the data to build a model, it is a commonly used method to prevent overfitting during training.

Each round of cross-validation involves randomly partitioning the original dataset into a training set and a testing set. The training set is then used to train a supervised learning algorithm and the testing set is used to evaluate its performance. This process is repeated several times and the average of accuracy and cross-validation error is used as a performance indicator.

Results

This chapter is divided into seven main sections, each one reporting the results related to one of the research questions. We have asked our subjects to judge the active character of each moral scene encoding a form of "affective" value, Perceived Affect (PA), which is directly dependent on the value of the circumstances embedding the actions. These results are based on experimental data obtained in the experiments 1, 2, 3 and 4. Data can be consulted in Annex 8.

From results obtained in experiment 1 (the Stereotype Affect; SA) and 4 (the Perceived Affect and the acceptability; PA and Acc), we have found that contextual modulation of moral decisions can be described by divisive normalization, an adaptive form of gain control that may thus be a general mechanism for sensory and cognitive computations.

From data collected in experiment 2 (the Perceived Affect; PA) we have showed that the affective value is independent from the acceptability.

Furthermore, experiment 3 (where the acceptability, Acc, was recorded) permitted to demonstrate that the cognitive load, eye scan paths, mouse trajectories, Reaction Times (RT) and Doubt Indexes (DI), were very different for the different types of moral responses.

And in addition, the values of Doubt Index (DI) allowed us to predict a subject's responses using a decoder and, further, to investigate how aware they were of the difficulty of their own moral decisions.

Finally, we have explored the difficulty of the dilemmas using results obtained in experiment 3.

Appendixes VI through XIV show in detail the different statistical analysis performed on the data. Moreover appendix XI shows a set of 5 special dilemmas whose particularity is the manipulation of the Perceived Affect (PA) scores in different directions in subsequent trials forcing participants to change of heart at least two times during the course of the dilemma.

1. Affect Based Gain Control by Divisive Normalization in Moral Judgments

Normalization is a canonical neural computation mediating divisive gain control and context dependent modulation in sensory and cognitive systems of the brain (Louie, Grattan, & Glimcher, 2011; Carandini, & Heeger, 2012; Louie, Khaw, & Glimcher, 2013; Louie, et al., 2014). Here we show that moral decisions depend on the perceived value (Affect) of the active character on a moral scene which depends on a comparable evaluation between its SA and the value of the circumstances in which the action is embedded (CA). It is hypothesized that, first, subjects perceive the values of SA and CA, second, a computation is made between the SA and the CA to obtain a subjective value, the PA, that explains linearly the moral decisions; values of PA that exceed 0 led to the moral acceptance of the dilemma, while values below that score to its rejection.

1.1 Stereotype Affect & Contextual Affect

In our everyday lives, one of the stimuli most likely to elicit affective responses is other people (Phelps, Lempert, & Sokol-Hessner, 2014). For most of the social tasks, like moral behavior, the shift in decisions may simple be due to the presence of another person.

The current analysis examines whether SA modulates moral Acc. SA scores for each target were averaged across participants from experiment 1 (Annex 8). We obtained the proportion of Acceptability for each trial from experiment 4 (Annex 8). A Spearman's correlation (figure 8) was run to determine the relationship between the average SA of the 10 active characters and Acc values of each moral dilemma. The findings confirm the correlation between both variables was rather weak (r_s= 0.2379; n = 40 (trials); p> 0.05) (figure 10).

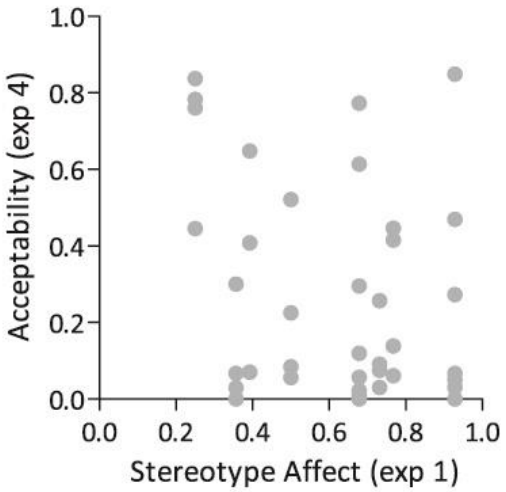


Figure 10. Stereotype Affect & Moral Acceptability. Scatterplot showing the correlation between SA average (exp 1) and Acc ratings (exp 4) [r_s= 0.2379; n = 40 (trials); p> 0.05]. For descriptive data please see annex 8.

As the data show, Acc ratings do not linearly depend on the SA. We believe that this inconsistency between the SA and the Acc may be due to choices not only depending on the qualities of the affective reaction evoked by the stereotype of the protagonist, but also on the contextual circumstances in which the action is embedded.

For instance, when people are asked to estimate the steepness of a hill, their estimates are influenced by the degree of effort they would have to make to climb the hill. Wearing a heavy backpack makes estimates higher; standing beside a friend makes them lower (Proffitt, 2006). These examples suggest that judgments are constructed on the fly and influenced by the context and by the feelings one has as one contemplates situations.

Thus, we modeled the influence of the context, proposing a Visual Analogy (figure 11), where the SA would be equivalent to the luminance of the center of a visual display (CL) and the CA would be equivalent to the luminance of the background (BL). The final perception of the central square, Perceived Affect (PA) in the case of the moral scenes and Perceived Luminance (PL) in the case of a visual scene, is a function of the contrast between the two variables.

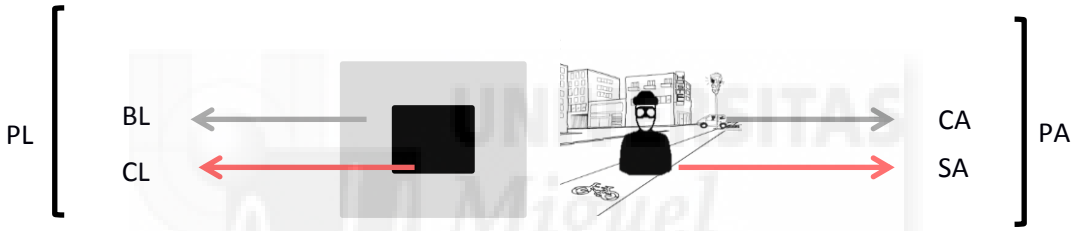


Figure 11. The Visual Analogy. Relationship between visual and moral parameters.

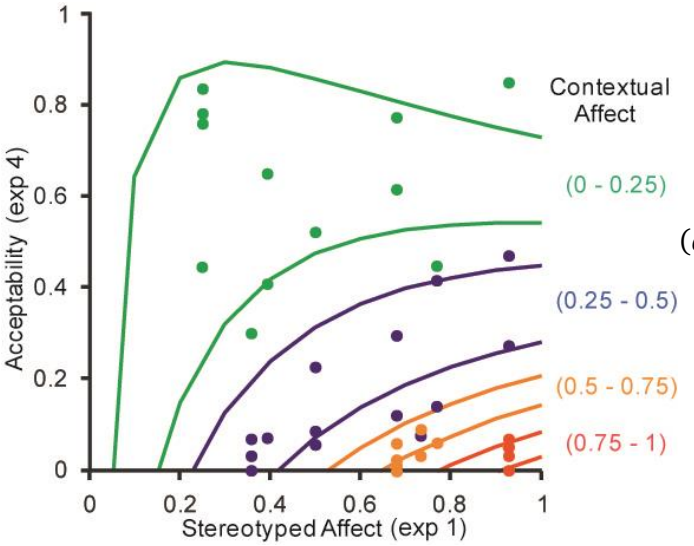
1.2 Affect Gain Control

Because DN is a computation repeated in a large number of systems, also in the visual system, we hypothesize that it should also underlie moral decisions (see Materials & Methods, section: 2.1). The CA scores were obtained indirectly by using the SA and the PA scores (figure 12). Coefficient values are shown in table 4.

Thus both the SA evoked by the protagonist himself and the CA evoked by the situational factors, are essential for moral decision making. Context impacts how people judge actions and actors in a way that is independent of purely deontological or absolute utilitarian considerations.

Table (4)
DN Coefficients (Figure 12)

Coefficient	**γ**	**σ**	**β**	**m**	**n**
Value	1.3	$7.95\ 10^{-2}$	0.6	1.25	1



$$(accep) = \gamma \frac{SA^n}{\sigma^m + (CA^m + SA^m)} - \beta$$

$$CA = SA\frac{(1 - PA)}{(1 + PA)}$$

Figure 12. Divisive Normalization Computation in Moral Decisions Making (1). Dots represent the relationship between Acc ratings (exp 4) and SA average (exp 1) n= 40 trials. Colour of dots reflects the empirical CA. Colored lines represent the theoretical values of Acc modeled by different contextual weight using divisive normalization equation which coefficients are presented in table 4.

1.3 Affect Contrast Gain Control

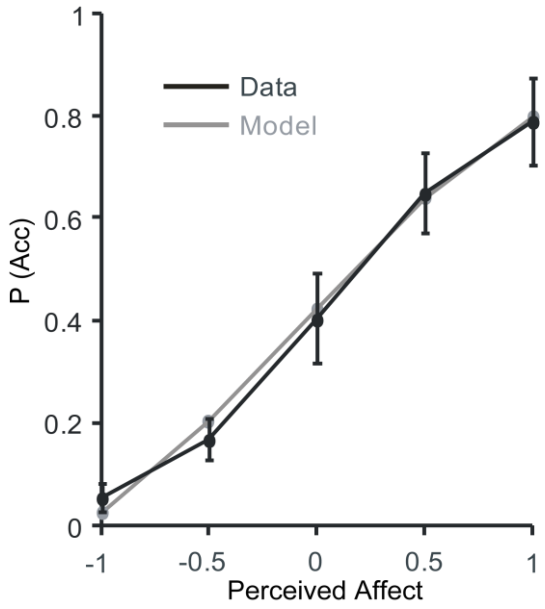
How is the affective reaction elicited by the protagonist in a given context incorporated into the computation of a moral decision? According to our Visual Analogy, moral decisions should be shaped by the same gain control mechanisms that operate in sensory pathways to represent value in a relative rather than absolute manner.

A follow-up analysis demonstrated that DN predicts moral acceptability. Affective processes elicited by the contrast between SA and CA influenced a factor underlying choice behavior: the computation of the PA, where larger values had higher probabilities of acceptance (figure 13). The values of the variables and coefficients of the DN equation that best fit the experimental data are shown in table 5. The divisive normalization model explained the empirical values of acceptability (R^2= 0.993; n = 5; p< 0.001; r_s= 1; n = 5; p< 0.05).

Appendix XIV shows the validation of the experimental data.

Table (5)
DN Parameters (Figure 13)

Coefficient	γ	σ	β	m	n	SA
Value	0.8750	$3.4725\ 10^{-6}$	0	1.1833	0.9142	0.9142



$$P(accep) = \gamma \frac{SA^n}{\sigma^m + (CA^m + SA^m)} - \beta$$

$$PA = \frac{(SA - CA)}{(SA + CA)}$$

Figure 13. Divisive Normalization Computation in Moral Decisions Making (2). Moral Acc as a function of SA and CA contrast. Black line represents experimental data n= 2572 trials. Grey line represents the modeled data obtained using the divisive normalization equation with values shown in table 5. The divisive normalization model explained the empirical values of acceptability [R^2= 0.993; n = 5; p< 0.001; r_s= 1; n = 5; p< 0.05].

1.4 Perceived Affect

The DN model linearizes the acceptability with the contrast between the SA and the CA, (or PA). The scatterplot in figure 14 shows the correlation between PA and Acc ratings (analyzed in each trial). The strong correlation suggests that PA has a large impact on the acceptability of a moral scene. (r_s= 0.9396; n = 40 (trials); p< 0.0001).

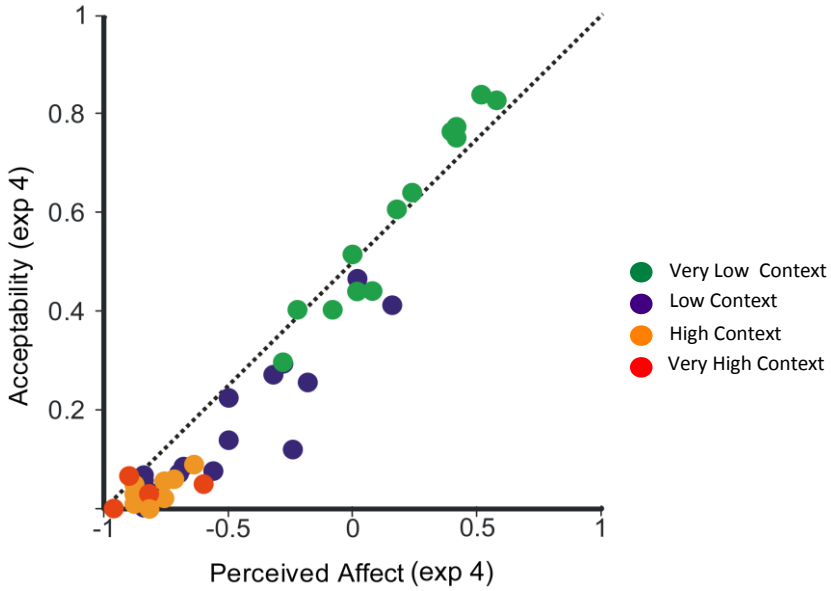


Figure 14. Divisive Normalization Computation in Moral Decisions Making (3). Scatterplot showing the correlation between the average of PA (exp 4) and the Acc ratings (exp 4). Colors indicate the CA interval value. The correlation was significant [r_s= 0.9396; n = 40 (trials); p< 0.0001].

In a more detailed analysis (appendix VII), we have examined the strength of the relationship between Acc and PA calculating χ^2. In this case, instead of using the average of both variables per trial (n= 40), we have employed every single performance of the participants, taking into account the dilemmas' composition (1-2-3-4 trials).

Table (6) shows the intervariable correlations among the different experiments. A few details warrant mentioning. First, all the data suggest a very strong positive correlation between PA and Acc. Second, this correlation was similar even when the data was recorded in different experiments and with different participants.

Table (6)
Intervariable Correlations (r_s)

	Acc Exp.3	**Acc Exp.4**
PA Exp. 2	0.8903	0.8606
PA Exp. 4	0.8461	0.9396

Table (7) shows the intravariable correlations among different experiments. As in the previous table, all the data suggest a very strong positive correlation revealing that in both experiments, 3 and 4, participants answered in the same way.

Table (7)
Intravariable Correlations (r_s)

	PA Exp.4	**Acc Exp.4**
PA Exp. 2	0.8592	--
Acc Exp. 3	--	0.8768

Note. All correlations (table 6 & 7) were significant at p< 0.000,1 n=40 (trials)

1.5 Non-linear Responses

We found an individual variability in baseline affective tendencies that altered moral decisions. Individual's affects dispositions varied depending on the behavioral cluster where participants were classified. The findings confirm dramatic differences in PA average between deontological and contextual responses (figure 15). The statistical analysis is shown in appendix VIII. While CR, independently of the cluster, exceeded the value of PA=0, DR remained below this score. We observed a notable difference between cluster C1 and the rest of Contextual clusters. The C2, C3 and C4 categories needed a graded input for crossing the threshold.

Data seems evidence that moral acceptability is a question of threshold to PA, where the different Contextual clusters show the same behavior, it means, all of them reach the same output, but in response to different parts of the dilemma (hence contexts) (see also figure 16).

If we focus our attention to the critical trial (the first CR) and the previous one (the last DR), a huge difference, between both trials, in terms of PA could be observed. The first CR and the previous DR were subjected to Mann Whitney test to determine whether there were significant statistical differences between both trials (figure 17). Table (8) indicates the statistics values for all 10 dilemmas.

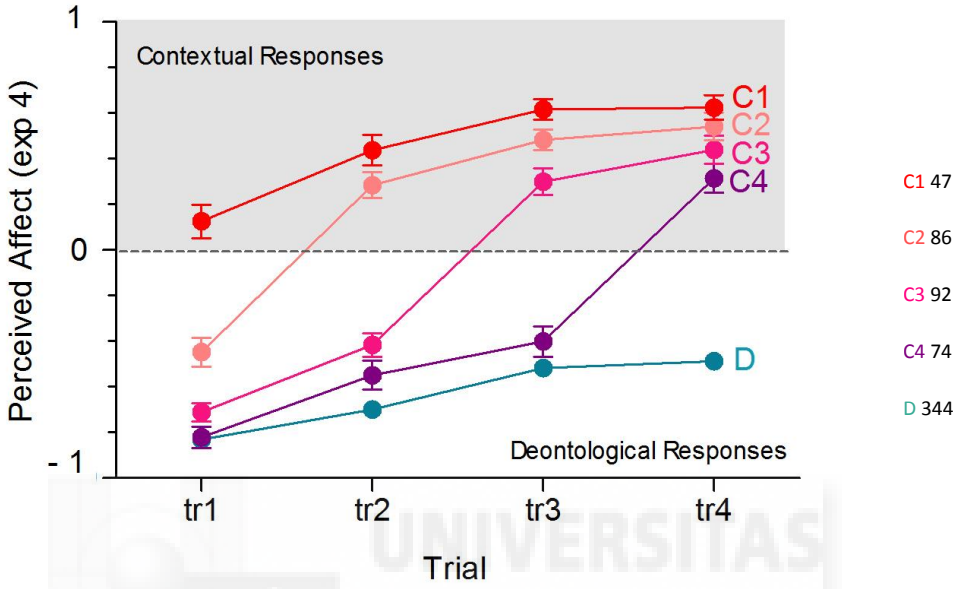


Figure 15. Divisive Normalization Computation in Moral Decisions Making (4). Relationship between PA and type of behavior. Each point indicates the PA average (± SEM) depending on the cluster and trial from experiment 4. Grey display represents the trials with a CR while white one represents the DR. The dashed line represents the threshold concept. Numbers on the right represent the sample's size of each cluster.

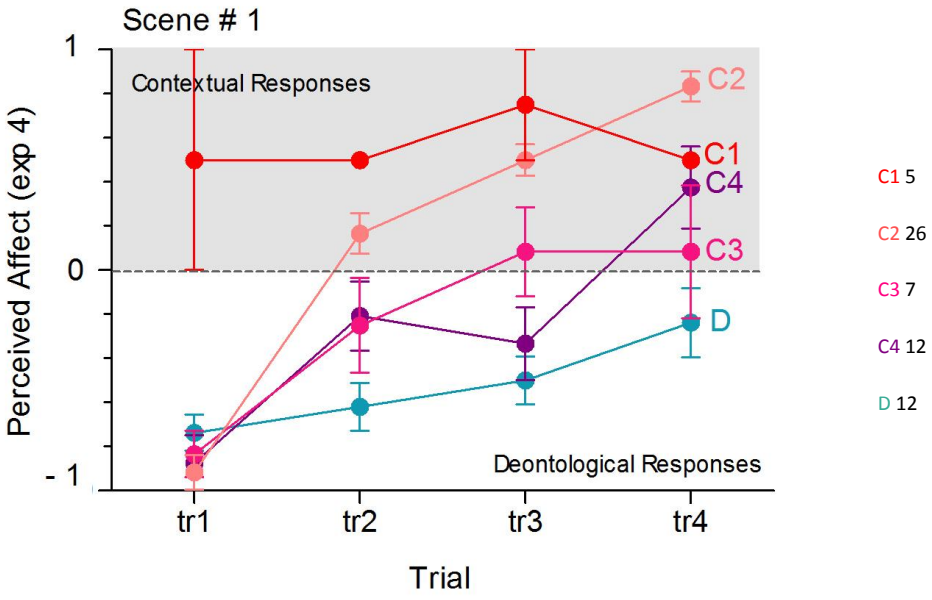
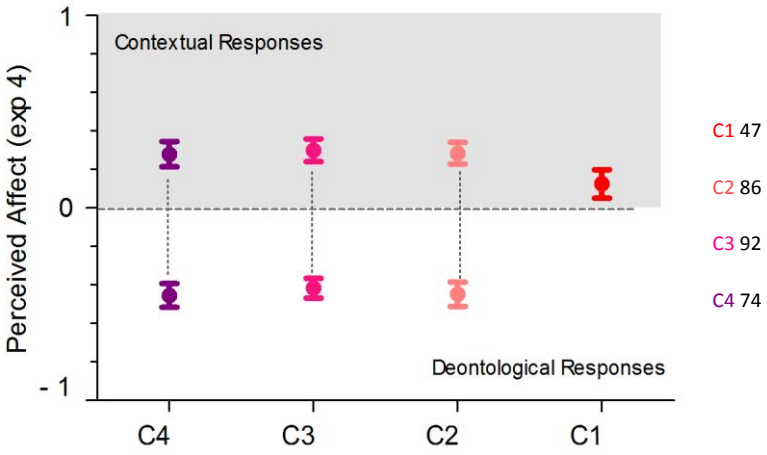


Figure 16. Divisive Normalization Computation in Moral Decisions Making (5). Relationship between PA and type of behavior in a single dilemma. Each point indicates the PA average (± SEM) depending on the cluster and trial from experiment 4. Grey display represents the trials with a CR while white one represents the DR. The dashed line represents the threshold concept. Numbers on the right represent the sample's size of each cluster.



Figue 17. Nonlinear Responses in Moral Decisions Making. Each point indicates the PA average (± SEM) made by all subjects who belong to the contextual's cluster for the first CR trial and the previous DR one from experiment 4. The horizontal dashed line represents the threshold concept. The vertical dashed line represents the transition between the last deontological response and the first contextual response. All comparations were significant at p< 0.0001. Numbers on the right represent the sample's size of each cluster.

Table (8)
DN Threshold

	C4	**C3**	**C2**
	Tr 4 vs Tr 3	Tr 3 vs Tr 2	Tr 2 vs Tr1
U	1333	1439	1348
n	74	92	86

Note that all comparisons were significant at p< 0.0001

With respect to contrast levels, which were calculated adapting the Michelson contrast formula (see Materials & Methods, section: 2.2), a few details, warrant mentioning (figure 18). Specifically, C1 differed significantly from the other clusters due to the large PA which then makes it impossible to show a high contrast. A possible explanation for this might be that the dynamic range was very low; meaning that the difference between the smallest and largest value of PA was small.

We have conducted a Kruskal-Wallis test to address how PA contrast scores differ between the different Clusters. Differences were found between C4 and C1 (K= 23.66; p< 0.001), C3 and C1 (K= 17.06; p< 0.05) and C2 and C1 (K= 20.35; p< 0.01).

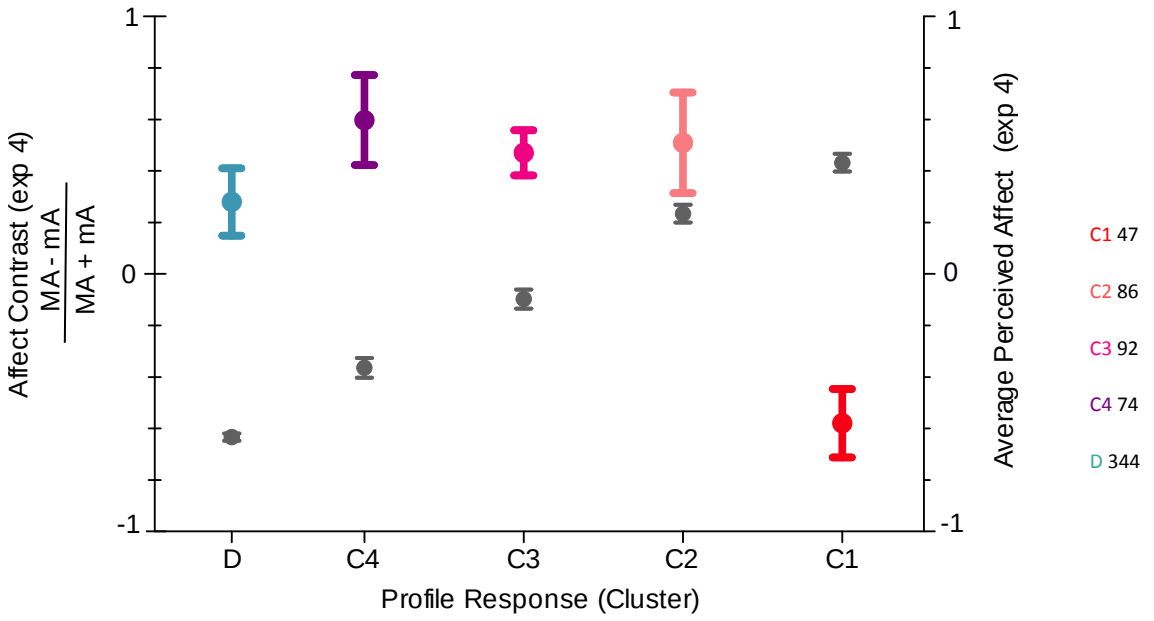


Figure 18. Perceived Affect Dynamic Range: Contrast Levels. Relationship between PA contrast and type of behavior from experiment 4. Hot colors represent the PA contrast values average (± SEM) for each cluster. Grey dots indicate the PA average (± SEM) of each cluster. Numbers on the right represent the sample's size of each cluster.

Figure 19 shows that C clusters do not differ significantly in maximum PA which is always above 0 (unlike D cluster). The key aspect that distinguishes them is the trial where flues reach threshold. Differences were found between C4 and C1 (K= -97.82; p< 0.05), C4 and C2 (K= -106.4; p< 0.01), D and C4 (K= - 155.2; p< 0.001), D and C3 (K= -216.8; p< 0.001), D and C2 (K= -261.6; p< 0.001), D and C1 (K= -253.1; p< 0.001).

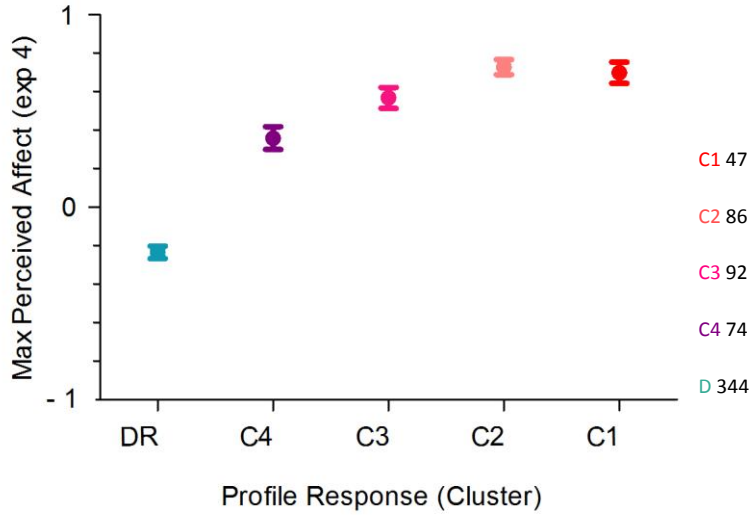


Figure 19. Perceived Affect Dynamic Range: Maximum Perception of PA. Hot colors represent the PA maximum level average (± SEM) for each cluster from experiment 4. Numbers on the right represent the sample's size of each cluster.

2. Acceptability & Perceived Affect

In this section we tried to determine, first, whether the participants had a specific style for responding moral dilemmas, i.e. to what extent they tend to respond following normative rules rather than other circumstantial considerations. Second, the type of behaviors that were found in experiments 3 and 4 and whether their relative distribution was similar in both experiments. Finally, to what extent men and women showed similar patterns of responses.

2.1 Acceptability Index

The previous section indicates that the perception of the moral acceptability of an action depends on the PA of its "protagonist". As we have shown, this perception is modulated by contextual non-moral information. Thus, we next evaluated how much each participant was influenced by context using an index (Acceptability Index, AI, see Materials & Methods, section: 2.3) that measures the percentage of contextual responses per participant.

Figure (20) shows the frequency of the different AI scores in our sample. The histogram shows a symmetrical distribution with very similar mean, median and mode, (M= 0.29; Median= 0.28; Mode= 0.2); and the Shapiro-Wilk normality test confirmed that the rate of acceptability is distributed normally (n=35; $r^2 = 0.8717$; w=0.8232; p> 0.05).

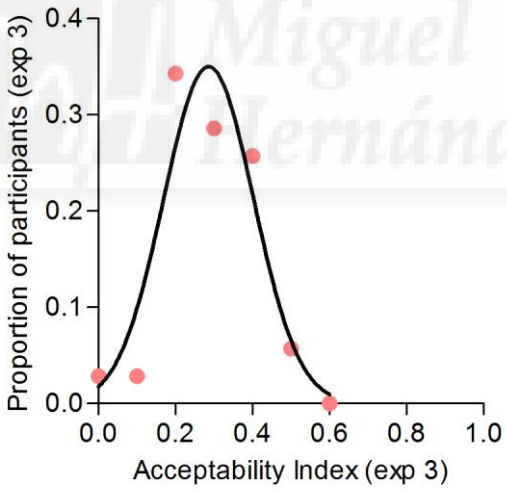


Figure 20. Acceptability Index Distribution. Histogram representing the frequency of AI scores. Shapiro-Wilk normality test confirms the Gaussian distribution [W= 0.9728; p> 0.05; $r^2 = 0.8911$; n= 35].

2.2 Similar Distribution of Acceptability in Experiments 3 and 4

Figure (21) shows the distributions of behavioral outcomes (clusters) in experiment 3 and 4. A χ^2 Test was performed to test the similarity of both distributions. Just the C2 cluster showed statistical differences between the two experiments (χ^2= 5.843; df= 1; p<0.05). Nonetheless, it is striking how similar the two distributions were, particularly if we take into

account that they were obtained in two different countries (Exp. 3 in England and Exp. 4 in Spain) and included different nationalities.

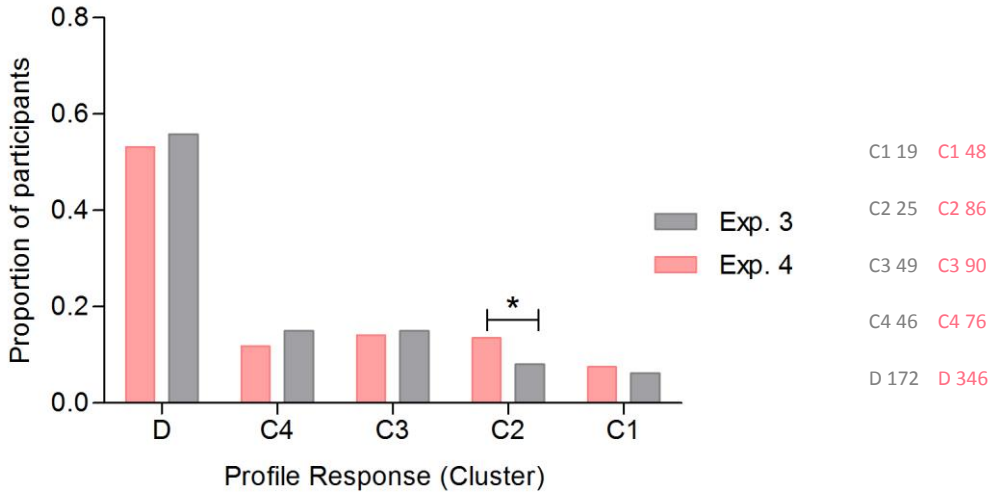


Figure 21. Distribution of Different Behavioral Outcomes. Histogram showing the proportion of the different clusters in experiments 3 and 4. A χ^2test revealed a statistical significant differences only in cluster C2 [X^2= 5.843; df= 1; p < 0.05]. Numbers on the right represent the sample's size of each cluster.

2.3 Men and Women Differences regarding Affect and Acceptability

We next performed an analysis to test whether there was also a difference in acceptability ratings and PA between men and women (figures 22, 23 & 24). As figure 21 shows, we have found differences between men and women in Acc ratings of experiment 3 (χ^2= 11.21; df= 1; p <0.001), however this difference did not exist in experiment 4 (χ^2= 1.637; df= 1; p= 0.2008).

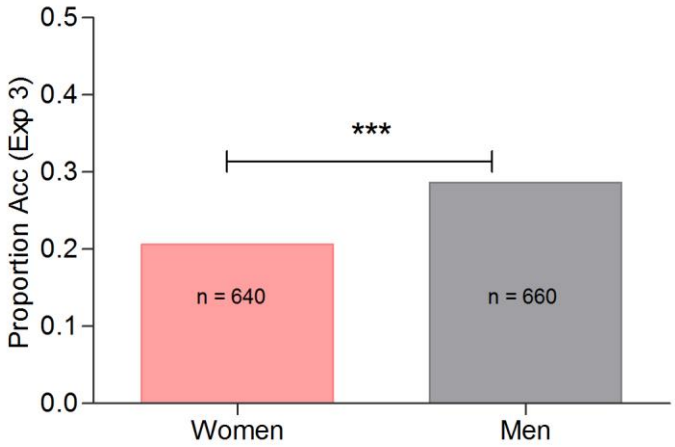


Figure 22. Acceptability Ratings by Women and Men in Experiment 3. The histogram shows the percentage of dilemmas that were considered morally acceptable by men and women. A Chi Square's exact test revealed statistical differences between women and men in experiment 3 [X^2= 11.21; df= 1; p< 0.001; n_{women}= 640 (trials) n_{men}= 660 (trials)].

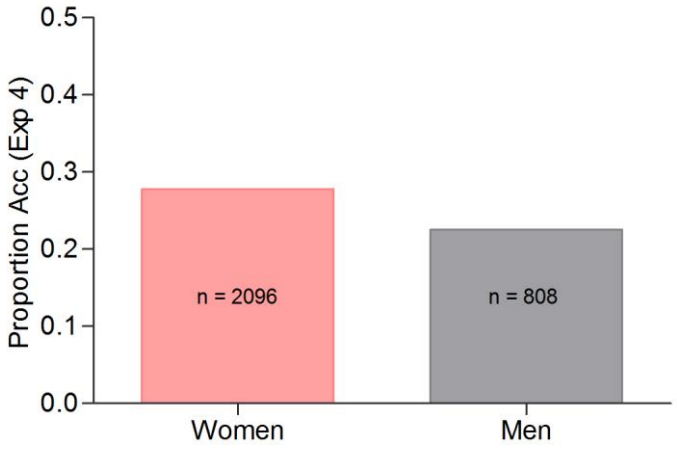


Figure 23. Acceptability Ratings by Women and Men in Experiment 4. The histogram shows the percentage of dilemmas that were considered morally acceptable by men and women. A Chi Square's exact test did not reveal statistical differences between women and men in experiment 4 [X^2= 1.637; df= 1; p> 0.05; n_{women}= 2096 (trials), n_{men}= 808 (trials)]

A Mann Whitney test showed there were statistically significant differences in PA between men and women (figure 24) (U= 799747; p< 0.05).

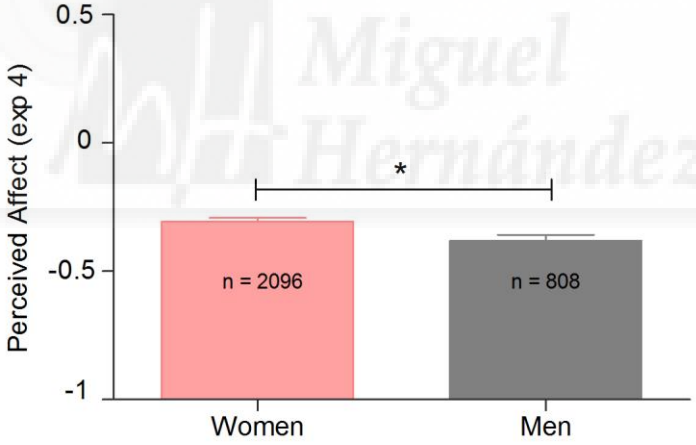


Figure 24. Perceived Affect Scores by Men and Women. Mean PA (± SEM) by men and women in experiment 4. A Mann Whitney test revealed statistical differences between women and men responses [U= 799747; p< 0.05; n_{women}= 2096 (trials), n_{men}= 808 (trials)].

3. The Timing of a Moral Judgment

In this section we compare the RTs associated to moral (with deontological and contextual answers) and control responses.

3.1 Different Latency between Moral and non-Moral Questions

We decided to compare the moral responses with the control answers to explore if the different RTs are associated to the type of stimuli (moral, visual or simple questions). First, we calculated the average of the 3 positive and the 3 negative control questions (see Materials & Methods, section: 1.3.2) for each subject. Then we normalized the rest of the participant's RTs (3 false controls, the 27 visual trials and the 40 moral items) to the average of the positive and negative controls. Figure (25) shows the normalized RTs in moral and non-moral responses of three participants. These subjects were selected because they obtained very different scores in the AI between them. As we can observe, normalized RTs of all type of responses tend to be equal. The control RTs' values are shown in table 9.

Table (9)
Control RTs' Descriptive Values

s. 25	M	SEM	n	s. 32	M	SEM	n	s. 27	M	SEM	n
	1.991	0.77	6		2.238	0.15	6		1.441	0.19	6

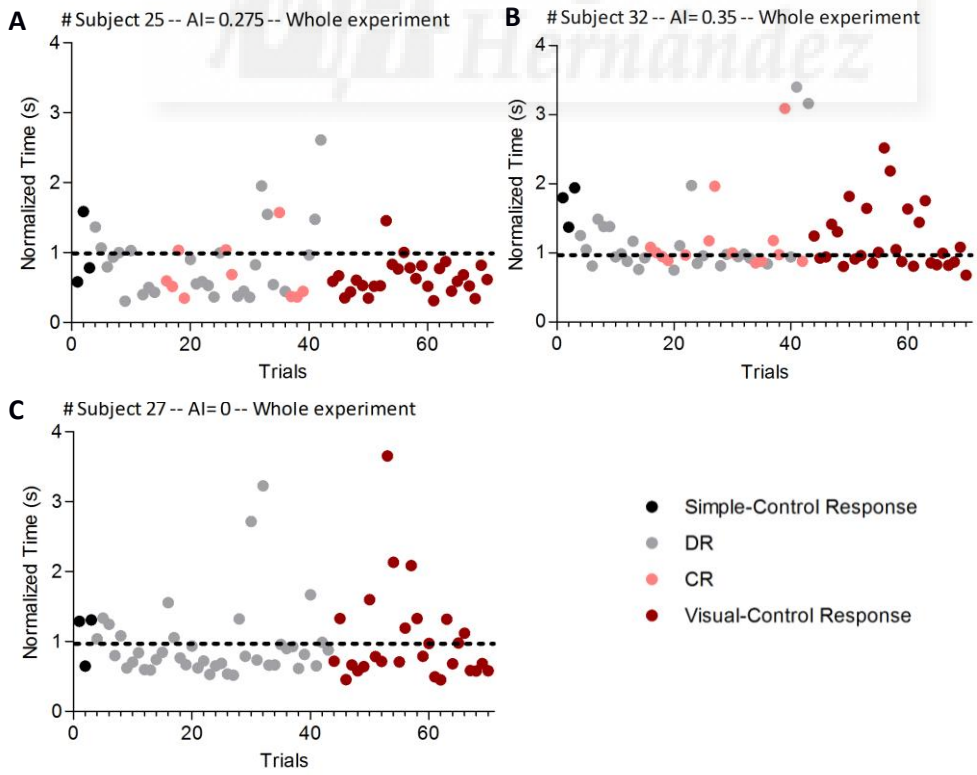


Figure 25. RTs Normalized at Control Questions. A) RT of all the responses made by a subject with a 0.275 AI. B) RT of all the responses made by a subject with a 0.35 AI. C) RT of all the responses made by a subject with a 0 AI. Dashed line represents the RT of the control questions.

3.2 The Timing of a Moral Judgment

In this section we compare the RTs associated to contextual and deontological response.

We constructed a Generalized Linear Model (GLM_z) to evaluate, which are the explanatory variables that led us to predict RTs depending on the type of response. We have used 1286 trials. In this case, the model used the following information:

- Dependent Variable (DV): RT
- Probability Distribution: Tweedie (1.5)
- Link Function: Log
- Independent Variables (IVs): Dilemma, Trial order in the dilemma (nested effect) and Acceptability.
    - Number of levels of the IVs:
      - Dilemma: 10
      - Order trial of the dilemma (nested effect): 4
      - Acceptability: 2

Table (10) shows that we have found a main effect of Dilemma and Trial order but not of Acceptability on RTs.

Table (10)
Test of Model Effects – Deontological & Contextual Trials

	Wald χ^2	df	p
Dilemma	100.868	9	**< 0.0001**
Order trial (Dilemma)	2117.1	30	**< 0.0001**
Acc	0.320	1	0.571

Table (11) shows that the predicted and experimentally recorded RTs are very similar (in terms of mean values).

Table (11)
Estimated and Real Mean Values – Deontological & Contextual Trials

	Estimated Values		Real Values		
	M	**SEM**	**M**	**SEM**	**n (trials)**
No Acc	2.916	0.267	2.955	0.101	969
Yes Acc	3.052	0.282	3.257	0.228	317

Empirical and predicted values were correlated r= 0.227; p< 0.01; n= 1286.

Empirical RT's of both contextual and deontological responses were compared using the Bonferroni's post hoc analyses. There were no significant differences between both types of responses (p= 0.572; df= 1)

Figure (26) shows the similarity between the average RTs (A) and distribution (B) of deontological (M= 2.955, SEM= 0.101) and contextual responses (M= 3.257, SEM= 0.228).

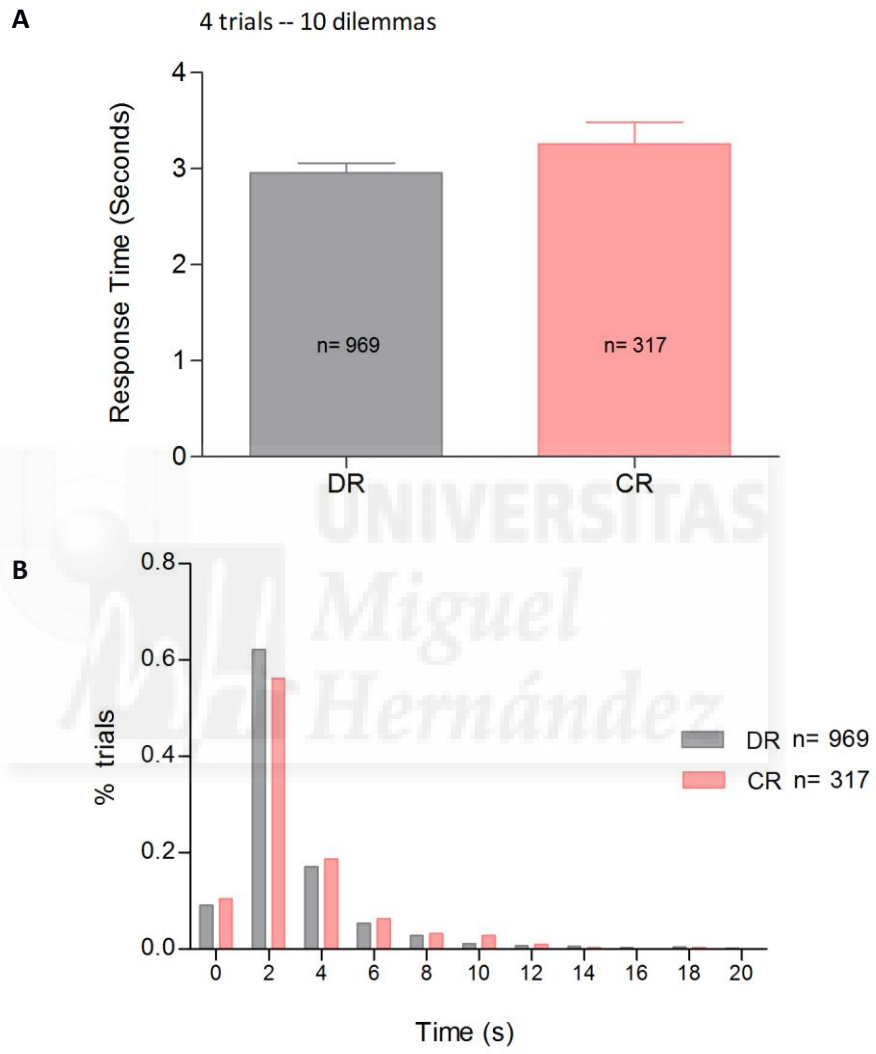


Figure 26. RT Differences between CR and DR. A) Mean RTs (± SEM) by response type in experiment 3. A post hoc bonferroni's test did not reveal statistical differences between deontological (n= 969) or contextual (n= 317) responses [p> 0.05]. B) Histogram showing the RTs distribution of deontological and contextual responses in the same experiment.

We have constructed our second GLM_2 to test specifically in those dilemmas where we have found both types of responses in the first trial, whether the Acceptability was a key variable in predicting RTs. We used 98 trials. The model used the following information:

- DV: RT
- Probability Distribution: Tweedie (1.5)
- Link Function: Log
- IVs: Dilemma and Acceptability.
  - Number of levels of the IVs:
    - Dilemma: 3
    - Acceptability: 2

Table (12) shows that there was no main effect of neither the Acceptability nor the Dilemmas.

Table (12)
Test of Model Effects – Deontological & Contextual Firsts' Trials

	Wald χ^2	**df**	**p**
Acc	0.320	1	0.195
Dilemma	100.868	2	0.371

Table (13) shows that the predicted and experimentally recorded RTs were also very similar in this case (in terms of mean values).

Table (13)
Estimated and Real Mean Values – Deontological & Contextual Firsts' Trials

	Estimated Values		**Real Values**		
	M	**SEM**	**M**	**SEM**	**n (trials)**
No Acc	3.747	0.319	3.707	0.532	74
Yes Acc	2.969	0.465	3.052	0.399	24

Empirical and predicted values were correlated r= 0.219; p> 0.05; n= 98. We have corroborated that this model is not a good one, probably because the number of trials is too small.

Empirical RT's of both types of responses were compared using the Bonferroni's post hoc analyses. There were no significant differences between the RTs of contextual and deontological trials (p= 0.172; df= 1) (figure 27).

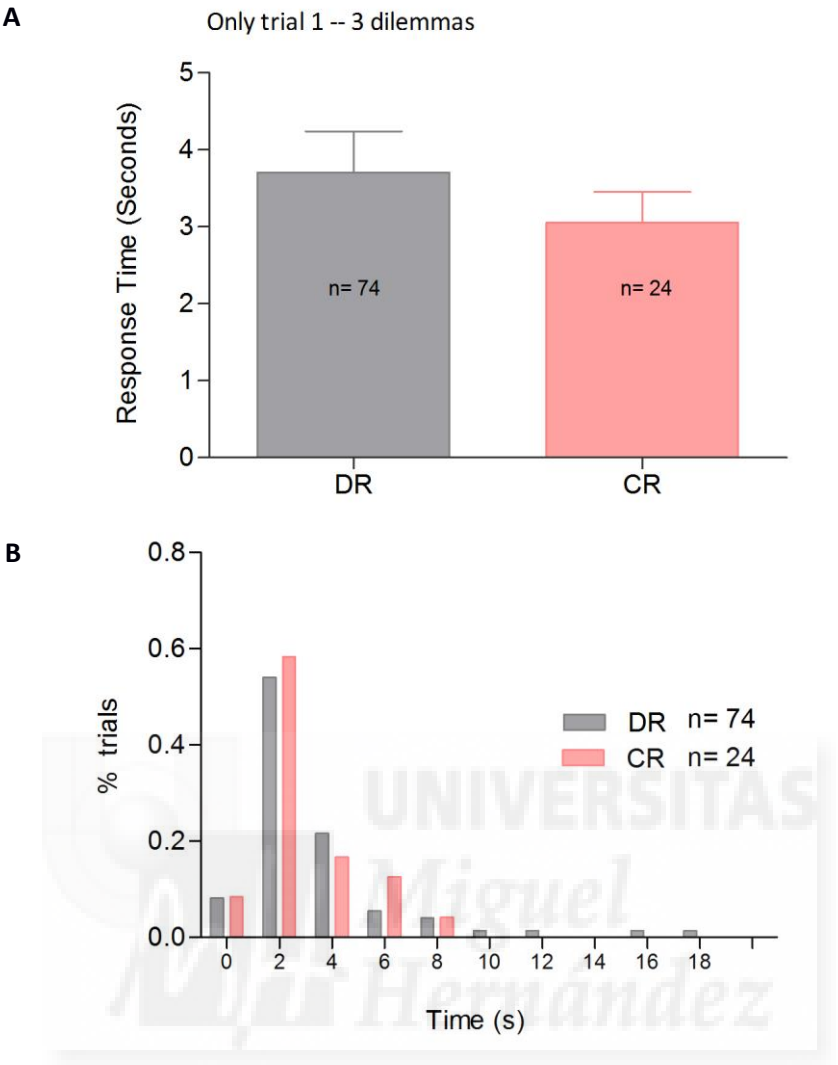


Figure 27. RT differences between CR and DR in the first trial. A) Mean RTs (± SEM) by response type in experiment 3 in dilemmas that had both CR and DR in the first trial. A post hoc bonferroni's test did not reveal statistical differences between deontological (n= 74) or contextual (n= 24) responses [p> 0.05]. B) Histogram showing the RTs distribution of deontological and contextual responses in dilemmas that had both types of responses in the first trial.

Finally, we have constructed our last GLM_z in this section, this time considering the trials where the acceptability ratings were above 5%, to evaluate whether the Acceptability was a key variable responsible of the RTs. We have used 873 trials. The model used the following information:

- DV: RT
- Probability Distribution: Tweedie (1.5)
- Link Function: Log
- IVs: Dilemma, Trial order in the dilemma (nested effect) and Acceptability > 5%.
- Number of levels of the IVs:
  - Dilemma: 10
  - Trial order in the dilemma (nested effect): 4
  - Acceptability > 5%: 2

Table (14) shows a significant main effect of the predictive variables Dilemma and Trial order in the Dilemma but not Acceptability.

Table (14)
Test of Model Effects – Deontological & Contextual Trials with more than 5% of Acc

	Wald χ^2	df	p
Dilemma	62.675	9	**< 0.0001**
Order trial (Dilemma)	58.223	17	**< 0.0001**
Acc >5%	0.590	30	0.442

Table (15) shows that the predicted and the experimentally recorded RTs were again very similar (in terms of mean values).

Table (15)
Estimated and Real Mean Values – Deontological & Contextual Trials with more than 5% of Acc

	Estimated Values		Real Values		
	M	**SEM**	**M**	**SEM**	**n (trials)**
No Acc	3.012	0.939	3.090	0.141	560
Yes Acc	3.139	0.132	3.252	0.230	313

Empirical and predicted values were correlated r= 0.236; p< 0.01; n= 873.

Empirical RT's of both types of responses were compared using the Bonferroni's post hoc analyses. There were no significant differences between contextual and deontological trials (p= 0.445; df= 1) figure (28).

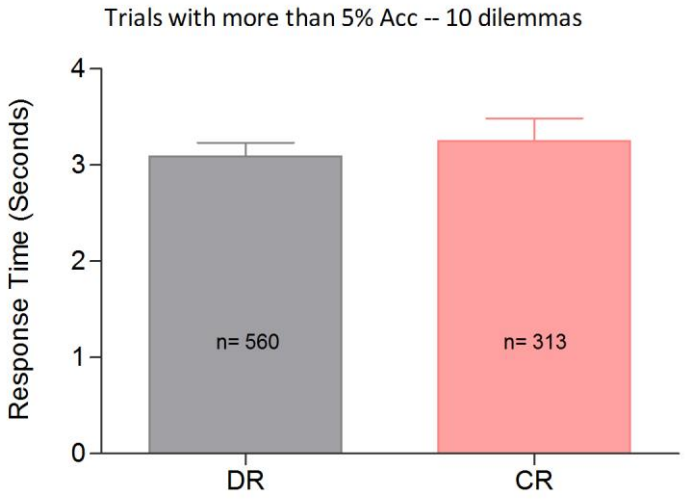


Figure 28. RT differences between CR and DR with more than 5% of acceptability ratings. Mean RTs (± SEM) by response type in experiment 3. A post hoc bonferroni's test did not reveal any statistical difference between deontological (n= 560) or contextual (n= 313) responses [p> 0.05].

Finally, we have compared the latency of CRs in the trial 4 of dilemmas that ended up with an uncertain outcome. Figure (29) shows that responses tended to be faster in fixed trials (average vs. average) probably because they incur in a lower cognitive load. The differences in RTs were significant (Mann-whitney (U) = 10014; p< 0.01).

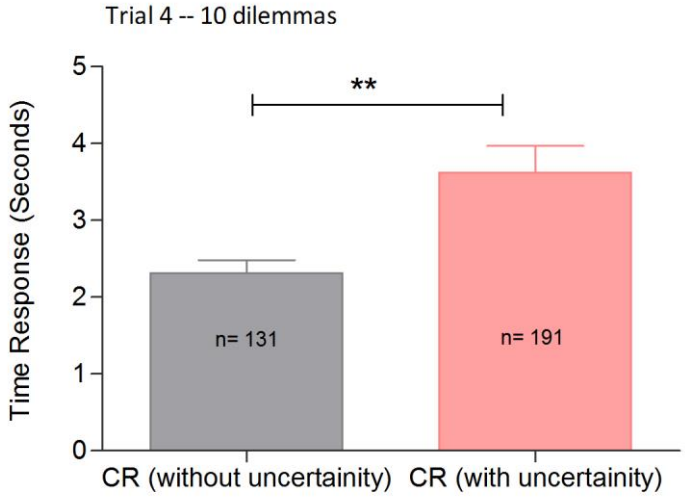


Figure 29. RT differences between trials with and without uncertainty. Mean RTs (± SEM) by trials with contextual response with or without uncertainty information in experiment 3. A Mann-Whitney test revealed statistical significant differences in latency between CR in trial without uncertainty responses (n= 131) and with (n= 191) [U= 10014; p< 0.01].

All the results shown in this section seem to suggest that there are not RTs differences between deontological and contextual trials.

4. Dynamic Aspect of Moral Judgments: Cognitive Loads Associated to a Moral Decision

This section of the chapter addresses the dynamic aspects of moral decisions. In particular, here we want to analyze if the cognitive load associated to a moral decision permits to predict the different response patterns (clusters) (Sweler, 1988).

First, in each run we identified two types of trials, swift and deliberative (see Materials & Methods, section: 3). Swift trials are those with the lowest scores for two of the three following parameters: RTs, # of visual fixations and changes in acceleration of the computer mouse on its way to the response box.

Our results show that there is, in fact, a specific distribution of swift and deliberate trials characteristic of each type of behavioral response (cluster). Results have revealed that there has been a specific pattern for each cluster. Figure 30 can be seen as a graphic description of each clusters' characteristic cognitive load.

- Deontological signature. This type of response is characterized by very fast responses, few fixations and a straight mouse trajectory in every trial. Average RTs were below 3 seconds in every trial. And, finally, there are no clear swift or deliberate trials.

- C4 signature. There is a gradual increase in cognitive load as the dilemmas proceed adding new context. Trial 4 is often classified as the deliberated trial, and trial 1 is usually the swiftest one. And, accordingly, trial 1 tends to be the fastest and trial 4 the slowest.

- C3 and C2 pattern. As in the previous case, there is a gradual increase of cognitive load as the dilemmas proceed adding new context and up until the first CR. The last DR and the first CR trials were the ones with the higher probability of being classified as deliberate, while the trials following the first CR were the swiftest ones. Average RTs also follow this pattern.

- C1 pattern. The deliberate trial was the first one in the run, while the swift trial was the last. Average RTs were below 3 seconds in every trial.

Taken together, these results suggest that all low cognitive loads are mainly associated with decisions that corroborate the acceptance of a previously evaluated moral scene, while high cognitive loads correlate with trials when decisions first departed from firm deontological rules or the immediately preceding trial.

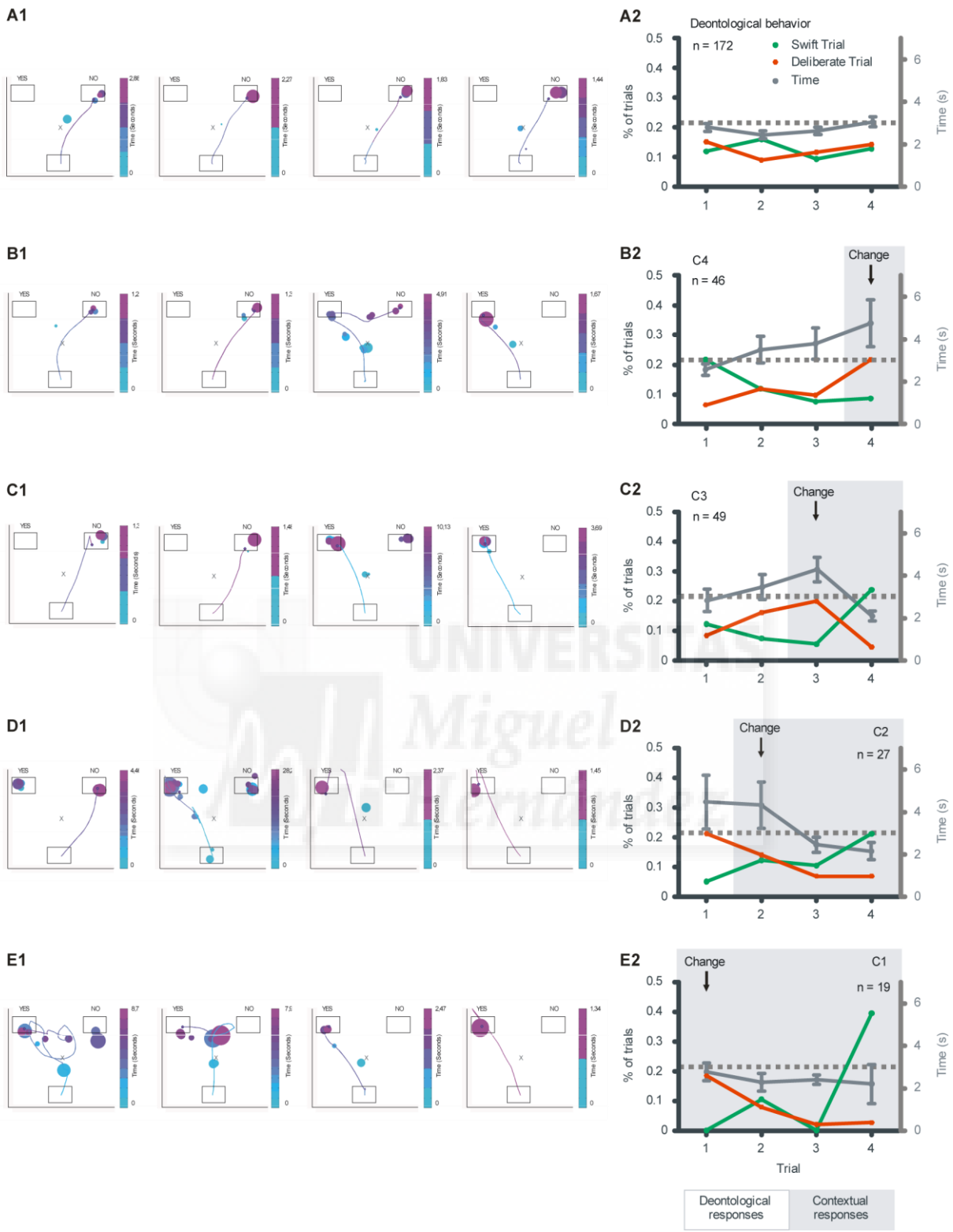


Figure 30. Dynamics of Response depending on Participants' Profile. A1-E1) Representative examples of eye scan paths and computer mouse trajectories for each cluster. A2-E2) Representation of trial's cognitive load based on the number of fixations, changes in mouse' acceleration and RTs. Grey line: Average Reaction Time (± SEM) to make deontological or contextual (shaded) decisions. Green line: proportion of trials with low cognitive load (swift). Red line: proportion of trials with high cognitive load (deliberate).

Note that the Mouse's Track time is correlated with the fixation time.

We have constructed a GLM_z to evaluate how different aspects of the moral dilemmas affected RTs; and further, whether we could use that information to predict similar RTs in each cluster. We have used 1234 trials. The model used the following information:

- DV: RT
- Probability Distribution: Tweedie (1.5)
- Link Function: Log
- IVs: Dilemma, Trial Order in the Dilemma (nested effect) and Cluster.
  - Number of levels of the IVs:
    - Dilemma: 10
    - Trial order in the dilemma (nested effect): 4
    - Cluster: 5

All the IVs proposed, showed a significant main effect on RTs Table (16).

Table (16)
Test of Model Effects – Clusters

	Wald χ^2	**df**	**p**
Dilemma	77.611	9	**< 0.0001**
Order trial (Dilemma)	2934.025	30	**< 0.0001**
Cluster	18.129	4	**< 0.001**

Table (17) shows that the predicted and experimentally recorded RTs are very similar (in terms of mean values).

Table (17)
Estimated and Real Mean Values - Clusters

	Estimated Values		**Real Values**		
	M	**SEM**	**M**	**SEM**	**n (trials)**
D	2.707	0.264	2.735	0.099	688
C4	3.581	0.450	3.763	0.378	184
C3	3.490	0.538	3.325	0.274	180
C2	3.046	0.414	3.324	0.412	108
C1	2.076	0.312	2.441	0.275	76

Empirical and predicted values were correlated r= 0.120; p< 0.01; n= 1234.

5. Predicting Moral Decisions

We used a decoder based on the Doubt Index (DI) (see Materials & Methods, section: 2.4 & 2.5) scores of each participant to investigate to what extent we could predict both the dilemma they were responding to, and their type of response (cluster they belong to).

5.1 Predicting a Moral Behavior (Cluster)

Table (18) and figure (31) show the prototypical DI (M ± SEM) for each one of the five possible clusters in each trial.

Table (18)
Prototypical ID

	Tr1		**Tr2**		**Tr3**		**Tr4**	
	M	SEM	M	SEM	M	SEM	M	SEM
D	0.432	0.041	0.352	0.037	0.436	0.031	0.411	0.04
C4	0.497	0.042	0.492	0.068	0.567	0.038	0.637	0.036
C3	0.504	0.034	0.580	0.04	0.582	0.05	0.360	0.065
C2	0.584	0.051	0.568	0.052	0.497	0.045	0.402	0.053
C1	0.458	0.13	0.405	0.089	0.517	0.088	0.202	0.122

Note. n= 33 subjects (participant 7 and 14 were taken out of the analysis due to technical problems with the eye Tracker recordings)

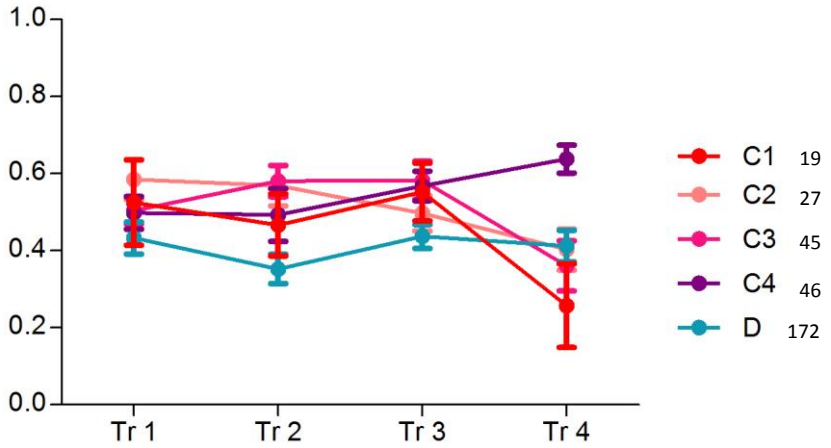


Figure 31. Doubt Index Distribution. DI average (± SEM) of experiments 3 in trials 1, 2, 3 and 4. Numbers on the right show the sizes' sample.

Decoding performance (0.3125 with p< 0.001) was significantly higher than chance (chance= 0.2). This system of prediction is very useful; we just need to know the DI of trials 3 and 4 (normalized at tr 1) to predict above chance the subject's behavior figure (32).

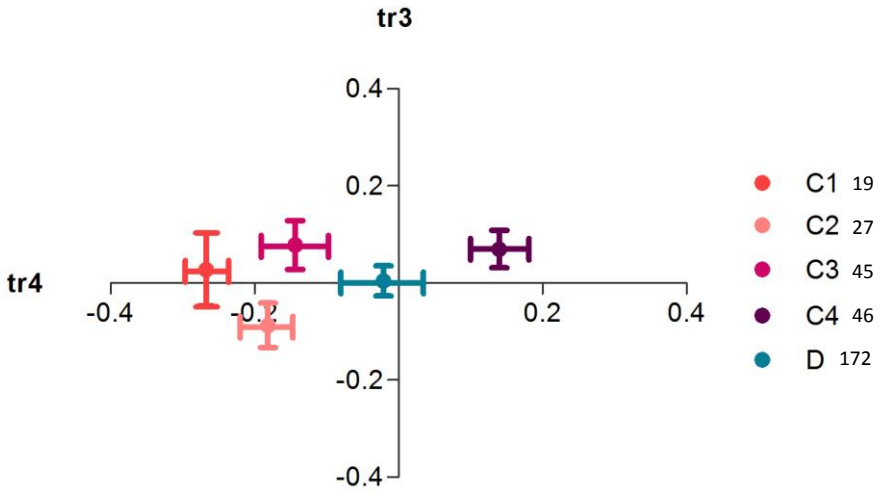


Figure 32. Fisher Linear Discrimination Decoder. DI average (± SEM) of experiments 3 in trials 3 and 4 (normalized at trial 1). Numbers on the right show the sizes' sample.

5.2 Predicting a Moral Dilemma

Table (19) shows the decoding performance of the 10 dilemmas. Chance here was the number of dilemmas (chance= 0.1). We could predict the subjects' responses in 6 out of the 10 dilemmas with p.value < 0.001. One could think that the dilemmas with the highest percentage of Deontological responses would be easier to predict, e.g. dilemmas 2, 5 and 9 (62.5%, 83.3% and 68.7% respectively). However this is not true for dilemmas 6, 8 and 10 (30.3%, 50%, 45.1%).

Table (19)
Prediction of Dilemmas using a Decoder

Dilemma	nhits/ntrials	p
1	0.2424	0.0141
2	**0.5333**	**<0.001**
3	0.2188	0.0358
4	0.2667	0.0078
5	**0.4444**	**<0.001**
6	**0.4375**	**<0.001**
7	0.20	0.0732
8	**0.4333**	**<0.001**
9	**0.3438**	**<0.001**
10	**0.3214**	**<0.001**

6. Awareness of the Difficulty of a Moral Decision

We have applied the central concept of signal detection theory to classify the performance of a binary classifier system, using the Doubt Index (DI) and the Confidence Report (CoR) (see Materials & Methods, section: 2.6 & 2.7).

Table (20) shows the four possible outcomes of the binary classifier:

Table (20)
Signal Detection Theory: Confusion Matrix of Experiment 3

	High DI	Low DI
Low CoR	143	54
%	22.9%	8.6%
High CoR	174	251
%	27.9%	40.3%

TP (True Positive) — FP (False Positive) — FN (False Negative) — TN (True Negative)

Data shows that 63.2% of the subjects were aware of their own deliberative process. We have performed an analysis to test how the awareness was distributed depending on the behavioral pattern.

6.1 Awareness of Moral Decisions by Cluster

We have elaborated contingency tables to determine whether a specific cluster in a specific trial had a particular behavior (TP, TN, FP or FN) compared to the overall behavioral possibilities in this specific trial and cluster. Table (21) shows the statistically significant results that have been found.

Table (21)
Awareness of Moral Decision by Cluster

Cluster	Type Behavior	trial	χ^2	df	p
D	FN	1	8.918	1	< 0.01
D	TP	1	9.590	1	< 0.01
D	TP	4	14.60	1	< 0.0001
D	TN	2	4.780	1	< 0.05
C4	FN	2	7.248	1	< 0.01
C4	TP	1	4.007	1	< 0.05
C4	TP	4	5.147	1	< 0.05
C3	FP	4	8.886	1	< 0.01
C3	TP	3	9.918	1	< 0.01
C3	TN	3	7.217	1	< 0.01
C2	FN	2	10.07	1	< 0.01
C1	FN	2	11.51	1	< 0.001
C1	TN	4	6.400	1	< 0.05

The awareness patterns are congruent with the cognitive load patterns' (see Result's section: 4). For instance, in the clusters D and C4, trials 1 and 4 were those where participants have correctly recognized a high level of doubt involved in the decision process. This is expected because, on the one hand, trial 1 is the most ambiguous due to its lack of contextual information. And, on the other hand, trial 4 is, for both clusters, the trial with the greatest percentage of deliberation. Moreover, Cluster C3 had in trial 3 the TP and TN. This is also expected because this cluster in trial 3 had the most deliberative trial or, on the contrary, for some subjects, the swiftest one. Cluster C1 had the TN in trial 4, this was also congruent, because is the easiest trail for this type of behavior.

In general terms, subjects have recognized the high DI values in those trials that were considered as the most deliberative ones. In addition the same phenomenon occurs with the identification of trials with low DI, coinciding with the trials identified as swift in the cognitive load pattern depending on the behavior. However, why people failed to recognize their own deliberative process in some instances is harder to explain and it can perhaps be related to post-hoc mechanism contributing to strengthen the subjects' confidence on their own decisions.

6.2 Signal Detection Theory Rates

We have also calculated the following general rates (table 22). We have used the TPR (proportion of hits) and the FPR (proportion of misses) to measure the d' (d'= 1.23). This represents the separation between the TPR (Hit) and the FPR (Miss) per participant (figure 33). Most of the participants have greater ratings in TPR than in FPR, suggesting that they could recognized the degree of doubt involved in their own moral decisions. In a more detailed analysis (appendix XIII), we have examined the ROC function for every single participant.

Table (22)
Signal Detection Theory Rates of Experiment 3

Rate	Proportion
Prevalence	0.50
Positive Predictive Rate	0.72
False Omission Rate	0.40
Accuracy	0.63
False Discovery Rate	0.27
Negative Predictive Value	0.52
True Positive Rate	0.45
False Negative Rate	0.54
True Negative Rate	0.82
False Positive Rate	0.17

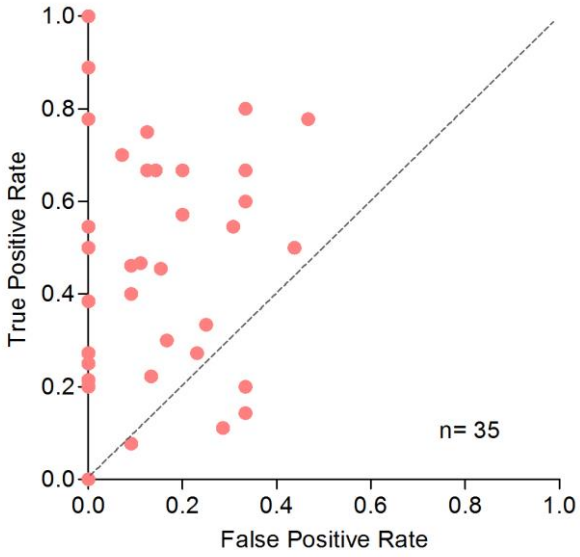


Figure 33. Awareness of Moral Judgments. ROC space graph showing the relationship between TPR and FPR per participant (n= 35) in experiment 3.

7. Difficulty of a Moral Judgment

This section has been included to analyze the relationship between the dilemma's difficulty and its acceptability. As was presented in the introductory section, most of the authors in the field propose that people's moral choices may shift according to the dilemma's difficulty (Greene, et al., 2001).

7.1 Difficulty based on Personal & Impersonal Dilemmas

According to Greene, some of the personal dilemmas prompt immediately deontological responses, but the most difficult evoke a competition between automatic and rational processes, and finally the quandary is responded in an utilitarian way. While, on the other hand, impersonal dilemmas are usually responded in an utilitarian way. To test this hypothesis, we have divided our dilemmas into these two sub-groups table (23).

Table (23)
Classification of Moral Dilemmas

	Dilemma	**Characteristics**
Personal	1; 3; 5; 7; 10	Action would cause: - serious body harm or death - to a particular person or group - the harm does not result from deflecting an existing threat
Impersonal	2; 4; 6; 8; 9	None of the above

We have correlated the PA of the active character in personal / impersonal dilemmas with the proportion of acceptability for each trial in experiment 4 (figure 34). The findings confirm that the correlation between the two variables was as strong in the personal quandaries (r_s= 0.8885; n = 20 (trials); p< 0.001) as in the impersonal dilemmas (r_s= 0.9707; n = 20 (trials); p< 0.0001). The data suggest people's moral decisions are driven by the subjective PA independently of the type of dilemma.

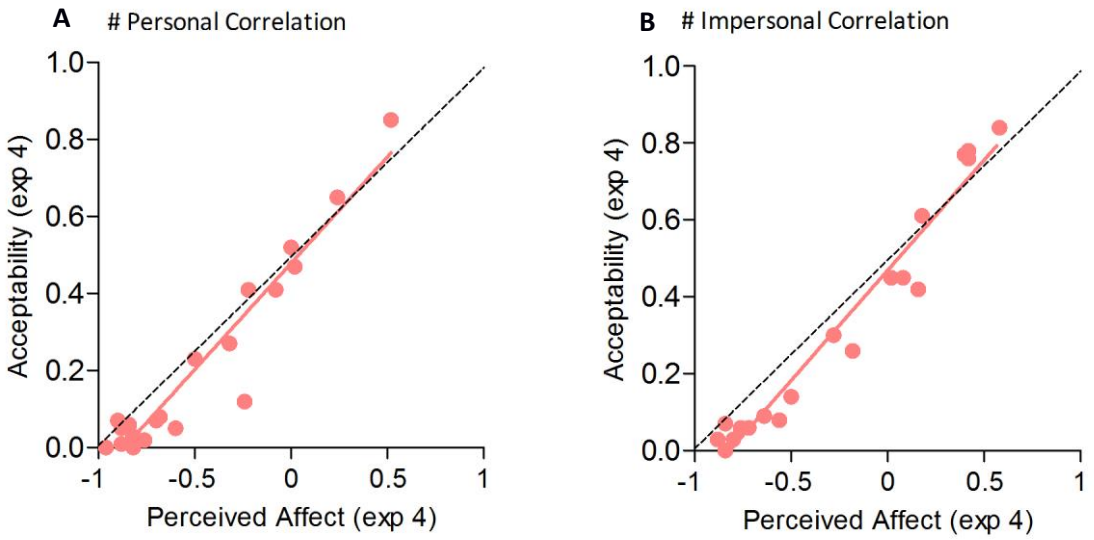


Figure 34. Personal & Impersonal Moral Dilemmas and the relation between PA & Moral Acceptability. A) Cross Correlation between PA average (exp 4) and Acc ratings (exp 4) in personal dilemmas [r_s= 0.8885; n = 20 (trials); p< 0.0001]. B) Cross correlation between PA average (exp 4) and Acc ratings (exp 4) in impersonal dilemmas [r_s= 0.9702; n = 20 (trials); p< 0.0001].

In this section, differences in terms of time between the different types of dilemmas, would be shown.

First, figure (35) shows the comparison of the latency between Deontological (M= 3.070; SEM= 0.153) and Contextual trials (M= 3.948; SEM= 0.521) in personal dilemmas. The differences in RTs were significant (U= 24961; p< 0.05).

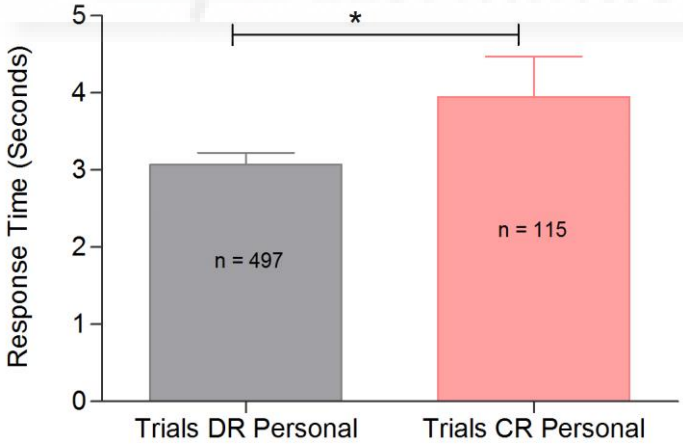


Figure 35. RT differences between CR and DR in Personal Dilemmas. Mean RTs (± SEM) by response type in experiment 3. A Mann-Whitney test revealed statistical differences between deontological (n= 497) and contextual (n= 115) responses [U= 24961; p> 0.05].

Second, figure (36) shows the comparison of the latency between deontological (M= 2.774; SEM= 0.132) and contextual trials (M= 2.712; SEM= 0.203) in impersonal dilemmas. There were no significant differences between both types of responses (U= 39606; p> 0.05).

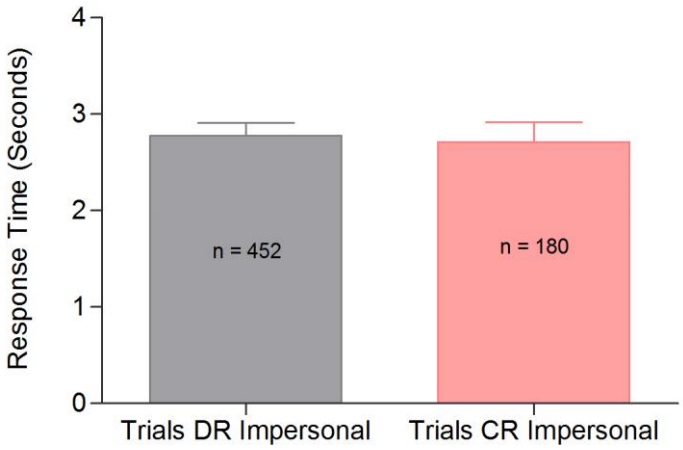


Figure 36. RT differences between CR and DR in Impersonal Dilemmas. Mean RTs (± SEM) by response type in experiment 3. A Mann-Whitney test did not reveal statistical differences between Deontological (n= 452) or Contextual (n= 180) responses [W= -1264; p> 0.05].

7.2 Difficulty based on Subject-Dilemma Interactions

We have used a Rasch model which assumes that the probability of a given person-item interaction is only governed by the difficulty of the item and the ability of the person.

The parameters of the model characterize the proficiency of the subjects and the difficulty of the items as locations on a continuous latent variable. In our experiment, items' scores represent the difficulty of the dilemmas. Thus, difficult dilemmas showed a high proportion of deontological responses. Accordingly, the subjects' choices represent the ability of the participants. High ability evaluations were obtained in those participants who responded in a contextual way in almost all dilemmas.

Figure 36A shows the Item Characteristic Curve (ICC) for our 10 dilemmas and also for hypothetical dilemmas with very high and very low difficulty. ICCs show the different probabilities to obtain a contextual response based on the participants' ability (marked with the gray band). The subject is likely to respond in a contextual way to the easiest dilemmas (with locations to the left and higher curves) and unlikely to respond "correctly" to difficult items (locations to the right and lower curves).

Figure 37B & 37C shows the Item ICC for the 4 trials which composed two different dilemmas. We can observe the difference between dilemma 1 and 7. While the first one has trials with diverse difficulty, the second one shares a very similar difficulty index between trials. The different fitting values are included in Appendix X.

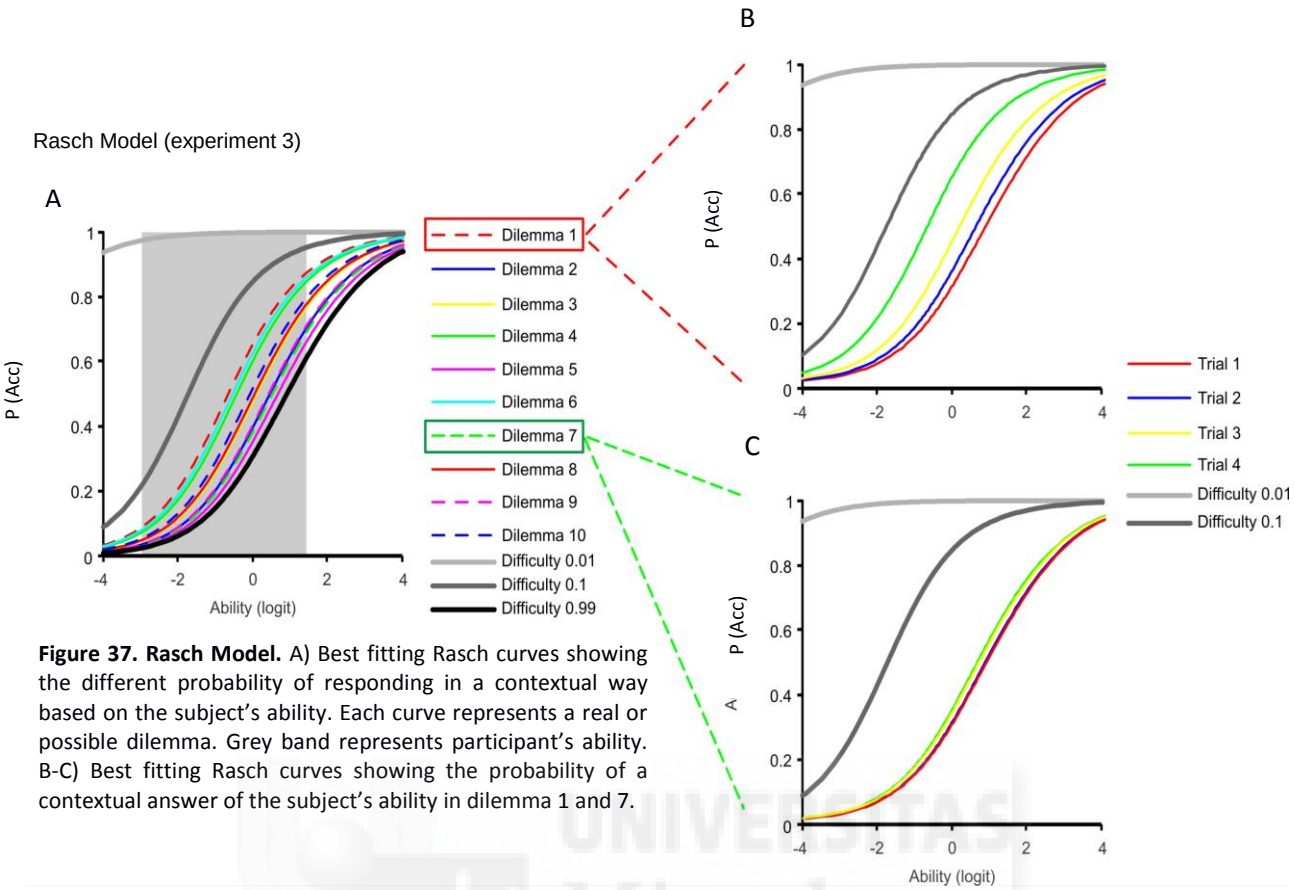


Figure 37. Rasch Model. A) Best fitting Rasch curves showing the different probability of responding in a contextual way based on the subject's ability. Each curve represents a real or possible dilemma. Grey band represents participant's ability. B-C) Best fitting Rasch curves showing the probability of a contextual answer of the subject's ability in dilemma 1 and 7.

Figure (38) shows how subjects responded the dilemmas in experiment 3 and 4.

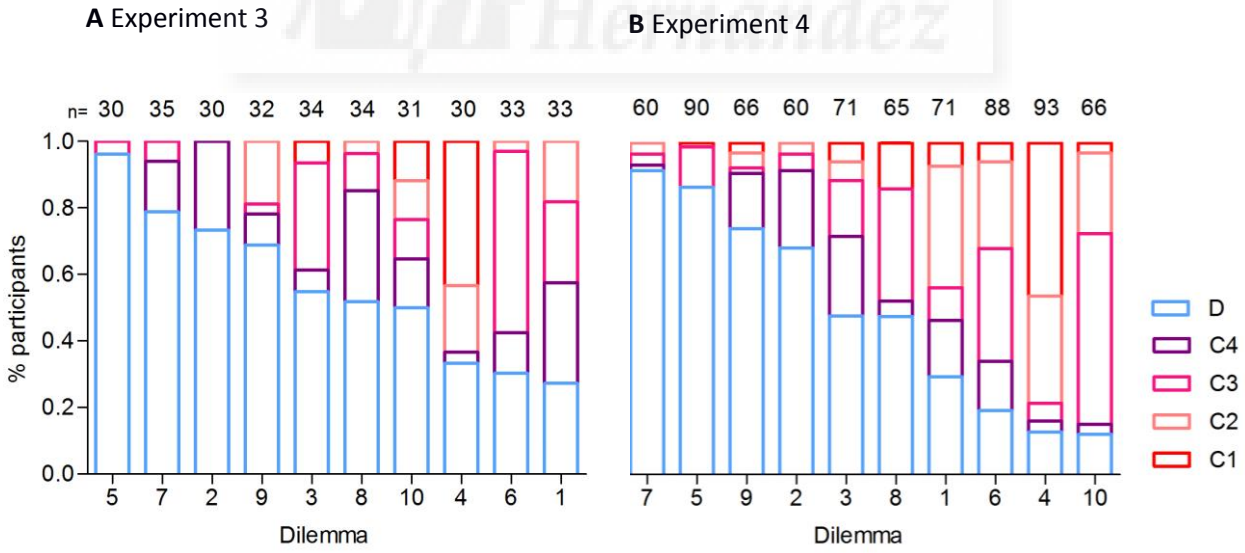


Figure 38. Behavioral Pattern per Dilemma. A) Bar graph showing the participants' behavior in experiment 3. **B)** Bar graph showing the participants' behavior in experiment 4.

Discussion

E.O. Wilson in 1975 predicted that ethics would soon be part of the "new synthesis" of sociobiology, in which distal mechanism (such as evolution), proximal mechanism (such as neural process) and the socially constructed web of meanings and institutions (as studied by the humanities and social science) would all be integrated into a full explanation of human morality (Wilson, 1975).

Our work materializes Wilson's consideration in an experimental way. We follow on the footsteps of previous approaches (Haidt, 2001) considering the conception of the "ingroup" (a social group to which a person psychologically identifies himself as a member) as the golden rule behind the neural mechanisms responsible for moral decisions. This mechanism is based on the perception of an affective value, which is active in the lives of people in diverse cultures, and is therefore influenced by social context and customs.

Specifically, we offer the first clear demonstration that affect (the subjective perception of affect and emotional aspects related to it) determines moral decisions, regardless of the nature of the dilemmas (killing person, cutting down trees, over boarding immigrants, bursting person, etc.).

And, contextual, situational (in any case, non moral) information is key to modulate Perceived Affect and hence moral decisions (It is also the first demonstration of these).

The goal of this section is to discuss the implications of these new findings, shifting the attention away from the DPM based on reasoning and emotions to intuitions and social factors.

Appendix XII shows a different analysis performed related to the RTs' and the rate of acceptability.

1. The Consensus of the Moral Field: the Role of Affect

Although the prevalent view of automatic processes is derived from the notion of a dual systems based on emotion and reason (Greene, et al., 2001), nowadays a new line of research is emerging which puts the emphasis on the role of affective perceptions in decision making. This work is based on how individuals' affective reactions may carry over to the assessment of subjective value that underlies the decision (Phelps, 2006; Keltner, & Lerner, 2010; Phelps, Lempert, & Sokol-Hessner, 2014).

Even if no grand unified theory of morality is supported, the major consensus is about the role taken by emotions and intuitions in this kind of choices (Damasio, 1994; Haidt, 2001; Greene, et al., 2001; Haidt, & Joseph, 2004; Greene, et al., 2004; Cushman, Young, & Greene, 2010). Nowadays the modal view in moral research is that reasoning and intuition both matter, but intuitions matter more. So we can say that it is assumed that Hume was mostly right.

In fact, we really experience countless moral dilemmas in our lifetime and frequently have the feeling that we resolve them quickly, unconsciously and without any apparent reflection upon explicit laws or utilitarian considerations, we decide about the acceptability of an action just following our intuitions (Haidt, 2001).

Moral intuition has been defined as the "sudden appearance in consciousness or at the fringe of consciousness, of an evaluative feeling (like-dislike, good-bad) about the character or actions of a person, without any awareness of having gone through steps of search, weighing evidence, or inferring a conclusion" (Haidt, & Bjorklund, 2008).

However, a more general concept was proposed, not exclusively related with moral choices. The "affect heuristic" is defined as a mental shortcut that allows people to make decisions and solve problems quickly and efficiently following the same like-dislike feeling (Slovic, et al., 2002, 2004). This heuristic is typically used while judging the risks and benefits of something, depending on the positive or negative feelings that people associate with a stimulus. It is the equivalent of "going with your gut". If the feelings towards an activity are positive, then people are more likely to judge the risks as low and the benefits high. On the other hand, if the feelings towards an activity are negative, they are more likely to perceive the risks as high and benefits low (Slovic, et al., 2004).

But which is the difference between the affective intuitions proposed by Haidt and Slovic's "affect heuristic"? In other words, are moral intuitions a special type of automatic processes? While Haidt describes moral intuition as a subclass of automatic processes, dedicated to discern between good-bad feelings (Haidt, 2001); the Slovic's heuristic applies to every type of stimuli/situations moral or non-moral (Slovic, et al., 2002, 2004).

Our data suggests that moral intuitions are not a specific type of perception, we constantly make like-dislike judgments, independently of the type of choice: moral, risk, aesthetics, etc. We propose that the major part of the events that we perceived, primarily those related with people, elicited an affect judgment, but this does not mean that there is a specific affective perception related with morality.

In experiment number 2, we asked our participants about the PA towards the active character on a dilemma without requiring them to make any moral evaluation. We correlated this PA data with participant's Acc data of experiment 3, obtaining a striking correlation r_s= 0.8903 (Figure 39). Our data thus give support to the emerging intuitionist theory: affective processes may influence a factor underlying choice behavior: the computation of a subjective value modulates behavior in a broad sense and, in particular, moral judgment (Osgood, 1962; Zajonc, 1980; Damasio, 1994; Phelps, Lempert, & Sokol-Hessner, 2014).

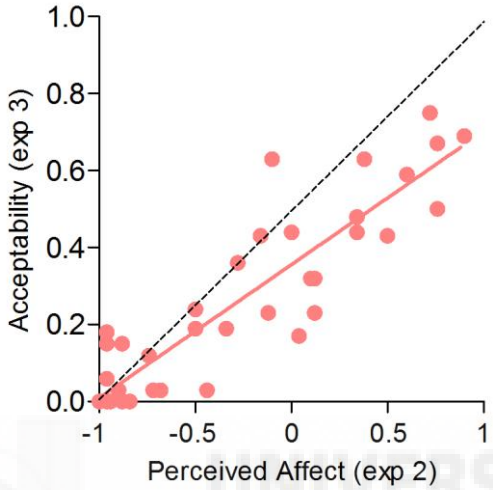


Figure 39. Affect Heuristic and Moral Judgment. Cross correlation between average PA (exp 2) and the Acc ratings (exp 3). The correlation was significant [r_s= 0.8903; n = 40 (trials); p< 0.0001].

This result does not necessarily imply that the intuitions triggered by moral situations, could be classified in Haidt's five cores: Harm/Care, Fairness/Reciprocity, Authority/Respect, Ingroup/Loyalty, Purity/Saucity (Haidt, & Joseph, 2004). In future research, it would be very interesting to classify and create dilemmas that specifically and independently address these five cores to evaluate the participant's acceptability profile in all these categories.

2. Affect Gain Control by Divisive Normalization

If PA is the final common pathway of morality, it is then fundamental to understand how this parameter is computed. To explore the neural mechanism behind moral decision making, we have established an analogy to the perception of luminance in the Visual System.

When we fixate on an object reflecting a specific luminance, say 200 (a.u.) embedded in a darker background with a luminance of, say 10 (a.u.), we categorize that central object as white. Under different lighting conditions, with a different surround luminance, however we could perceive it differently. If we translate this process into moral terms, we have two key elements, the affect evoked by the protagonist of a moral dilemma and the influence of the context in which the dilemma takes place.

The mechanism that we propose is not essentially different from the neural computation used by the visual system to perceive the luminance of an object. Like viewing an object under different lighting conditions, different values of PA are evoked depending upon

the conditions in which a scene has happened, in other words, depending on who is the protagonist of the dilemma (SA) and the influence of the context in which the action takes place (CA). In this way, the PA provides a guidance for what is morally acceptable or not (figures 13 & 14). In a Sherringtonian view, PA could be the final common pathway of morality. All contextual influences, past experiences or situational factors, will ultimately act modulating the PA.

The implementation of relative value coding via divisive normalization is in this case, functionally distinct from the classical DPM personal vs impersonal distinction. The computations performed are centered on individuals and not on the dilemmas themselves. Secondly, all the dilemmas are explained with the same mechanism using a single rule, independently of the type of dilemma. So, as Classical Normative Ethics, this new approach puts the emphasis on parsimony but focusing on the concrete and particular of the person, rather than the abstract and universal.

Divisive normalization was initially proposed to explain a number of nonlinear sensory responses, including cross-orientation suppression, surround suppression, contrast and saturation (Carandini, & Heeger, 2012). However, It has recently been shown to be fundamental for both sensory and non-sensory decisions (Louie, Grattan, & Glimcher, 2011; Carandini, & Heeger, 2012; Louie, Khaw, & Glimcher, 2013; Louie, et al., 2014). Our current results extend these observations to moral decisions and suggest that it could be a canonical mechanism for explaining all brain computations.

Current research on decision making has been transformed by the identification of different broad classes of algorithms that structure learning and behavioral choice, specifically tasks related with two alternative force choice (TAFC). Despite of there are many types of computational families, we have centered the discussion on the Sequential Samplings Models.

- Accumulator Models: In these models, evidence is accumulated separately for each possible outcome. This has the advantage that if many outcomes are possible, the models are simply extended by addition of one more variable for each additional alternative, with evidence for each alternative accumulating within its allotted variable (Smith, & Vickers, 1988).

- The Drift Diffusion Model (DDM): has been shown to describe accuracy and reaction times in human data for TAFC. The DDM assumes that decisions are made by a noisy process that accumulates information over time from a starting point toward one of two criteria or boundaries. The rate of accumulation of information is called the drift rate and it is determined by the quality of the information extracted from the stimulus (Gold, & Shadlen, 2007; Rangel, Camerer, & Montague, 2008; Ratcliff & Mackoon, 2008).

  It would be extremely interesting to model the RTs of contextual and deontological responses and the error rate (number of contextual responses obtained) using a DDM to evaluate whether with only one value of drift rate, we could be explain the RTs and the rate of both type of responses (see Appendix XII).

- The Leaky Competing Accumulator Model (LCA): The model integrates evidence for the two variables in support of one of the two alternatives in a TAFC task. The model includes a "leak" term, which forces a decay back to baseline for each individual variable in the absence of incoming evidence. The "competition" in LCA is a cross-inhibition between the two variables, thus improving upon original accumulator models in allowing evidence for one variable contributing to a reduction in the other variable (Usher, & McClelland, 2001).

- Neural Network Models and Attractor States: Models such as those respond to the match between a current item and memories of previously encoded items, which are contained within the network as attractor states. The network's activity reaches a stable attractor state more rapidly if the match is close (Hopfield, 1982).

All the algorithms presented here are centered on two different aspects. The first one is the information related with the past experience; the second one is the value associated to the stimuli at this precise moment. DN is a fusion between all these models since it considers the internal representation associated with each stimuli and the value of each scene in a particular state. We agree with Crockett (2013, 2016) and Kahane (2013) who consider that any successful algorithm for describing and predicting any decision requires the interaction of both sources of data.

Since our findings are consistent with an algorithmic implementation of divisive normalization in affective value coding, more research is needed to unravel where in the brain is the neural circuit that generates this representation.

By applying a specific algorithmic model to value coding, our results have equated the moral domain to any physical parameter our brain can measure. Morality would be like other physical phenomena, a system we can describe with laws because it exhibits law like regularities. We, as far as I know, are the first ones to present such a detailed mechanistic explanation for moral decision making.

Research in computational cognitive modeling, explores the essence of cognition. It embodies descriptions of neural mechanism in computer algorithms and programs, based on computer science (Turing, 1950). That is, it imputes computational processes onto cognitive functions, and thereby it produces runnable computational models. Right from the beginning of the formal establishment of cognitive science around the late 1970's, computational modeling has been a mainstay of cognitive science. Our results, could thus, inspire a new generation of moral automates for the new development in self driving cars, drones, and for long distant communications.

3. The Golden Rule of Morality: Affect Binds the Ingroup

Morality tends to be thought of as a group of codes and norms that should be respected by all the members of a group. Popularly, two types of codes are considered responsible of our moral behavior: deontology and religion. Religion has become central to discussion of morality. Across cultures and eras, people have often thought of religion as the

foundation of morality. Religions differ enormously around the world, but despite their diversity, for many people, moral wrongs are equated with religions wrongs, acts that violate the commandment of God.

The world's major religions generally include a well-developed set of practices and beliefs in order to suppress the differences between people. A very interesting study about these "moralizing" religions (Budhism, Jainism, Brahmanism, Daoism, Judaism, Stoicism, Christianity, Manichaeism and Islam) demonstrated that they emerged simultaneously in different parts of the world when human societies got so complicated probably to reduce the conflict between different groups (Baumard,& Boyer, 2013; Baumard, Hyafil, Morris, & Boyer, et al., 2015). Moreover the common rule between them is: "one should treat others as one would like others to treat oneself" (Neusner, & Chilton, 2008).

Our view is that this "golden rule" is the simplest expression of suppressing the differences in PA, according to which one should estimate the same affect for everybody in each situation. This means that you should favor your own group, but also the members of the outgroup that belong to your same society (religion).

But, in fact, people strive to protect their own social groups at the expense of outgroups to which they do not belong (Uhlmann, Pizarro, Tannenbaum, & Ditto, 2009). Moreover, individuals also exert more pressure on ingroup members. According to the "black sheep effect", people are generally less tolerant toward an ingroup member who transgresses social norms (Abrams, Rutland, & Cameron, 2003). In this situation, this particular individual would become a member of the outgroup, evidencing the dynamic character of the group pertinence. Group identity engenders ingroup favoritism, which in turn reinforces the boundaries between social categories favoring "us" versus "them" (Tajfel, & Turner, 1986; Brewer, 1999).

The capacity to discern "us" from "them" is fundamental in the human brain. Prejudicing and stereotyping stems from a mechanism of survival that structures the physical world (Amodio, 2014). When we confronted people in any social interaction, we unconsciously activate stereotypes that subsequently influence, unconsciously as well, our behavior. This has been shown with the Stereotype Content Model (SCM; Fiske, Cuddy, Glick, & Xu, 2002), a universal model which organizes group stereotypes along two social dimensions: competence (reflects traits that are related to perceived ability, intelligence, skill, creativity and efficacy) and warmth (captures traits that are related to friendliness, helpfulness, sincerity, trustworthiness, etc.) (Cuddy, Fiske, & Glick, 2008).

The SCM postulates that social groups fit within each of the four combinations of high and low levels of warmth and competence depicting one societal ingroup and three kinds of outgroups, predicting distinct behaviors and emotions. Admired groups judged as both warm and competent (the ingroup) elicit the desire to assist them and associate with them (active and passive facilitation). By contrast, contempt groups which are regarded as low in both warmth and competence are harmed passively or actively, that is, they are ignored or attacked. Pity people (low competence and high warmth), for example, may sometimes be neglected and other times patronized and offered too much help. Envied groups judged as competent and low warmth elicit active harm and passive association.

The basic dimensions of warmth and competence account for 82% of the variance in perceptions of everyday social behaviors (Wojciszke, Bazinska, & Jaworski, 1998). Subsequent data demonstrated that the SCM framework remains intact across cultures, predicting how groups are likely stereotyped, based on structural relations with other groups in their society, which is the behavior and emotion that those stereotypes elicit (Cudy, Fiske, Kwan, Glick, Demoulin, Leyens, & Ziegler, 2009).

Following the predictions of the SCM (Cuddy, Fiske, & Glick, 2008), an interesting study (Cikara, Farnsworth, Harris, & Fiske, 2010) was designed to ask whether moral acceptability depends on the stereotype associated with the person being sacrificed and the people being saved. In other words, whether participants may apply different rules to determine moral acceptability, depending on who is sacrificed and who is saved. The SCM hypothesis posited that saving high-warmth, high-competence targets should be the most morally acceptable because these people represent the ingroup. However, the results showed that neither warmth nor competence were the best predictors for explaining the acceptability ratings of the four groups.

In our view, Cikara, et al., (2010) have not considered the dynamic character of the group. People identify with and discriminate in favor of groups based on arbitrary and active criteria depending on the context (Tajfel, & Turner, 1986). We are like chameleons, designed to try out different colors to match our social partners. For instance, when we leave our home country and emigrate to another one, we began to establish a strong relationship with other compatriot, establishing a strong ingroup identity regardless of their particular personalities and even though in some cases we would have never had a friendship with some of them in a different context. Even more extreme, when we get into our car, the ingroup is restricted to yourself, while the rest of the drivers are immediately identified as the outgroup members.

We think that the phenomenon of grupal identity is closely related to the PA value. We suggest that once the PA is greater than a specific value, we start to consider this person as a member of our ingroup (figure 15) and we tend to find her actions morally acceptable.

4. The Dual Process Model vs Affective Model

A central aim of current research in moral Psychology is to characterize a workable dual system framework. The idea that opposing forces of emotion and reason compete in the human mind is prevalent for decision making in western thought. However, as we have mentioned above, this prevalent idea is starting to weaken in favor of an emerging view based on the important role played by intuitive affective reactions (Phelps, Lempert, & Sokol-Hessner, 2014).

Our results seem consistent with this affective approach applied to the field of moral decisions. In the next few sections we argue in favor of an affect gain control model and center the discussion around 3 of the main tenets of DPM: the classification of classical dilemmas, the different timing of deontological and utilitarian responses, and their different underlying neural mechanism.

4.1 Classical vs Ecological Dilemmas

Current research on human morality has largely focused on hypothetical moral vignettes, stripped of all non-essential contextual information (Foot, 1967; Thomson, 1985; Greene, et al., 2001; Moore, Clark, & Kane, 2008). There are some advantages of using these kinds of artificial moral dilemmas. First, by asking the individuals in these unfamiliar scenarios, researchers try to be more likely to guarantee impartiality in an effort to build a universally valid moral theory (Kahane, 2013). Second, employing artificial dilemmas is an easy way to compare all of them with each other. This practice allows researchers to look for the more relevant parameters (how the dilemmas should be formulated; how the dilemma should be conceptualized; and how should the relation between the participant and the dilemma be established) (Chistensen, Flexas, Calabrese, Gut, & Gomila, 2014).

However, stories about war, genocide, crimes against humanity and also, everyday life situations about hurting, stealing and lying, fill both the history books and the newspapers. In almost all conflict situations, we can observe how different people adopt different postures, most of the times absolutely opposite, some of them accepting as moral the actions that others consider totally unacceptable. This suggests a low congruence between dilemmas, showing the difficulty related with the idea of the existence of a Universal Moral theory. And also highlighting the complexity of identifying and classifying the parameters that influence a decision.

People's real moral choices drastically contradict the responses generated by the simple hypothetical probes. The underspecified and impoverished nature of the hypothetical moral dilemmas is unable to capture the complex social, emotional and motivational pressures inherent to real moral decisions (FeldmanHall, Mobbs, Evans, Hiscox, Navrady, & Dalgleish, 2012). Moreover philosophical dilemmas such as the Trolley (Foot, 1967) and the Footbridge (Thomson, 1985) scenarios, and also the battery elaborated by Greene and colleagues, capture a particular kind of moral tension, where the choice to harm another is offset against the greater good (Greene, et al., 2001). However our data shows that utilitarian considerations not always play a main role determining the response in moral quandaries. For evidencing this phenomenon we have introduced this dilemma:

1. You are an Algerian citizen who decides to travel to Spain aboard a dinghy. You and your travel partner push eleven compatriots overboard without anybody else noticing in order to ensure you get to your destination safely.
2. You decide to push them overboard when you are a few hundred meters from the coast of Spain, just as the dinghy begins to sink. You do not know how to swim.
3. You are one of 48 people on the boat. If all 48 of you stay on the dinghy, it will end up sinking and only the good swimmers will survive.
4. You need to find a job in Spain in order to help your family in Algeria survive.

As our results show, both, the stereotype of an Algerian immigrant and the situational factors that contextualize the action, influence the subjective value, the PA, underlying the moral judgment (figures 13 & 14). The number of people who are suffering the actions and

consequences of the action are simply two more pieces of information contributing to modify the PA, therefore, affecting the moral response (figure 40). However, this does not mean that trials (or dilemmas) with this type of contextual information, engage a different neural mechanism to make a decision.

Both, the traditional dilemmas and our new material, despite their differences regarding the importance of the situational factors embedding the scenes, have something in common: personal past experience elicited by the dilemmas has a key role influencing the moral choice.

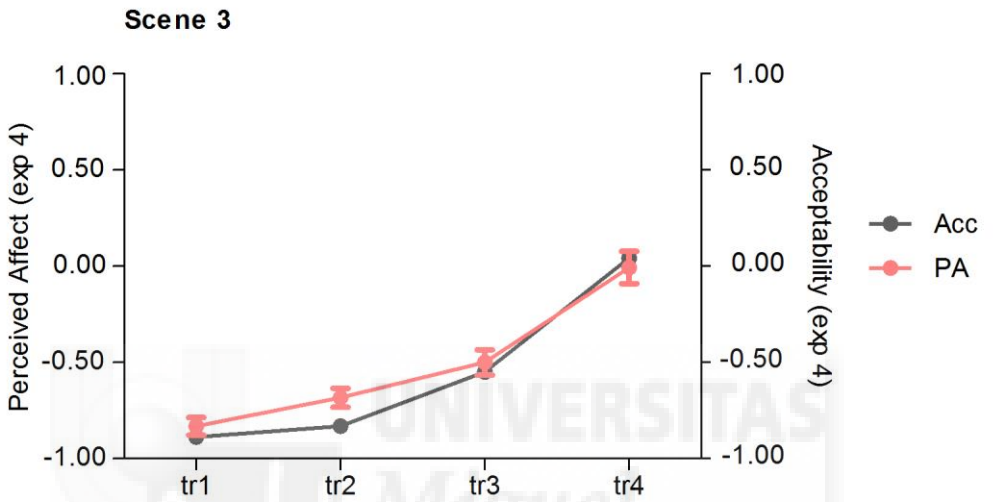


Figure 40. Relation between PA and Acc. Relation between the average PA (exp 4) and the Acc ratings (exp 4) n= 40 (trials).

When we evaluate whether an action is fair, we often consider our first-person experiences, and therefore, we decide whether something is moral by remembering both, the current subjective value and what we felt in the past in similar situations (Crockett, 2013). Basic decision making algorithms attend to this feature, as we have just seen in the previous section. We consider this type of information an inherent part of the context that has a large effect in how we process any stimuli, in a broad sense, and also in all types of moral dilemmas.

For example, vision scientist use visual illusions to explore what we see and what we potentially can see. Consider the illustration bellow, this figure is the demonstration that our personal experience contributes to make sensory judgments (figure 41).

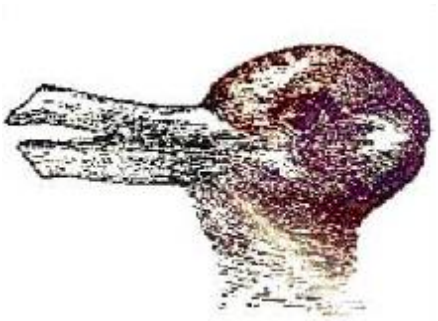


Figure 41. **Visual Illusion Rabbit-Duck.**

This representation is the earliest known version of the Rabbit-Duck illusion, an unattributed drawing of Fliegende Blätter, a German humor magazine. Some of us see a rabbit, others a duck and others, both, duck and rabbit. Interestingly, children tested on Easter Sunday are more likely to see the figure as a rabbit, whereas when tested on a Sunday in October, they tend to see it as a duck (Brugger, 1999). Our vital experience is partly responsible for labeling the figure as duck or as rabbit as it is also likely to be partially responsible for our perception of an action as right or wrong. Other interesting example is presented in Appendix V.

We have used a particular dilemma, one related with the Spanish terrorist group ETA, to highlight how contextual factors can influence choices:

1. A man decides not to give away a member of the ETA terrorist group, knowing that he was planning an attack.
2. The terrorist begged him not to inform the Police.
3. The member of the terrorist group was his own son.
4. Not giving him away means that the likelihood of a future terrorist attack occurring increases by 60%.

The majority of the participants considered morally unacceptable all the trials of the moral dilemma; and their responses were very fast, with very few fixations and with a very straight Mouse trajectory. However, a participant from the Basque-Country, home of ETA, performed the experiment with a very different pattern in comparison with the rest of the subjects (figure 42). It seems that her personal experience had a huge weight in the decision.

Greene considers his new rendition of the classical dilemmas as being like the drosophila is for biologists, as a little creature that a researcher can take into the lab and study in a controlled way, just as geneticists have learned a lot about human genetics by studying fruit flies (Greene, 2008). On the contrary, we think that these simple stereotyped dilemmas still suffer from the same influence of personal context and stereotyping as more ecological dilemmas such as ours do; therefore casting a doubt on their utility. Context is always present and thus extremely important in any type of decisions, and of course in the moral domain, because it is crucial for the computation of the PA. If context is not given to us we fill it in with our previous experience and knowledge about the structure of our social environment.

Nowadays, it is very difficult to assess to what extent the participant's personal experience is contributing to the decision. Future research must address specifically this issue.

Another interesting point regarding classical scenes is the classification of these dilemmas as moral personal or moral impersonal. The original empirical foundation of the DPM proposes a different kind of neural process depending on the type of dilemma (Greene, et al., 2001). Personal dilemmas are those that, require the harming (or killing) of another person or persons by an agent to achieve some goal, specifically when that harm is not simply redirected from one person or group onto another (that is, the agent must generate the harm themselves). When we classified our dilemmas following the general standard (personal vs impersonal) we have found that we can explain both kinds of dilemmas using the same

mechanism: participants used the PA as the factor responsible for the choice (figures 33, 34 & 35).

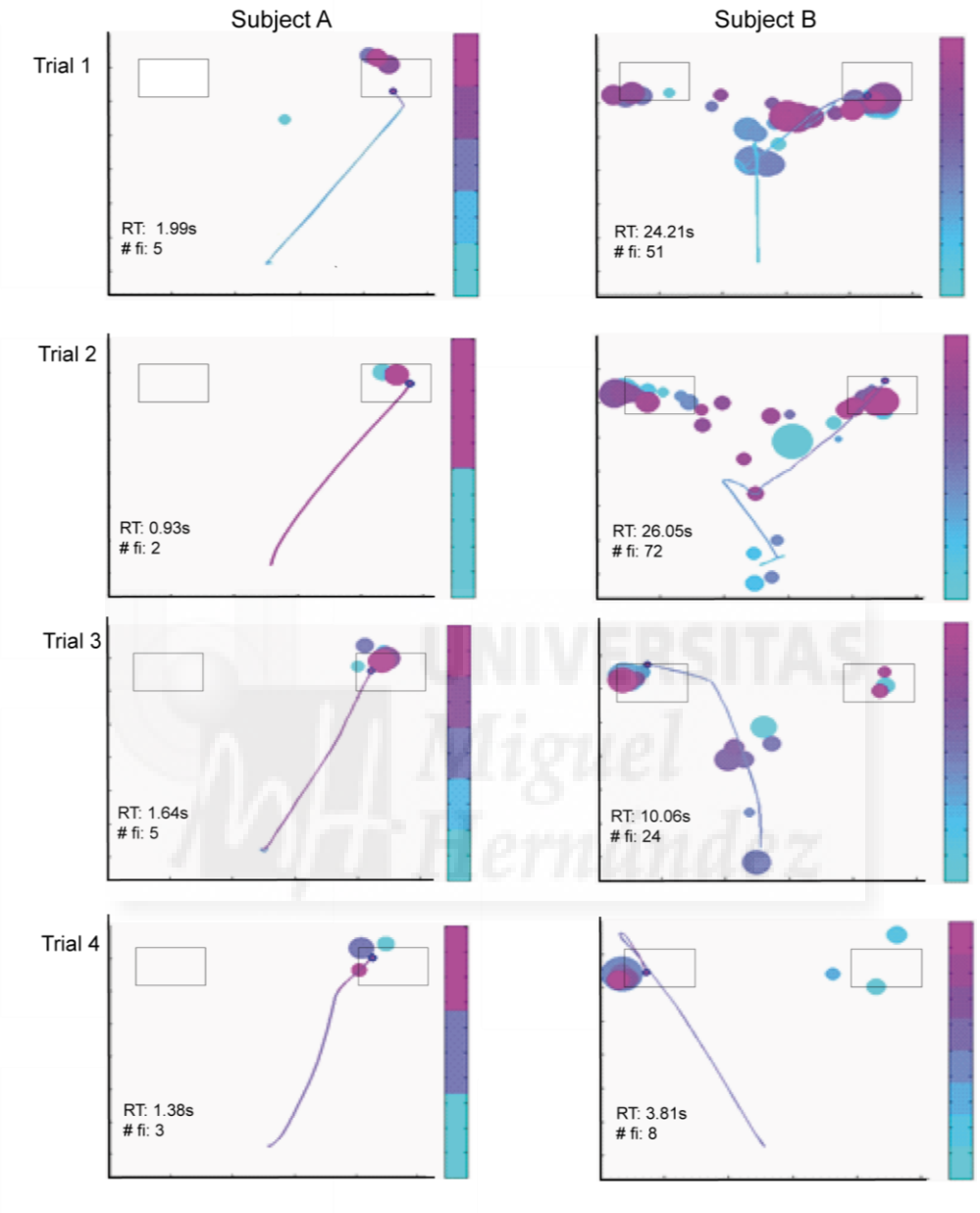


Figure 42. How Personal Context Influences the Dynamics of a Moral Decision. The section A shows the typical pattern found in this dilemma: Fast responses (RT), few fixations (fi) and a straight pattern in the mouse's trajectory. Section B shows the performance of one from the Basque-country, home of the Spanish Terrorist Group ETA.

Moreover some of these personal dilemmas are considered difficult to affront, for this reason the participants' responses are divided into deontological and utilitarian (Greene, et al., 2001). Impersonal dilemmas are those that do not satisfy at least one of the requirements for a personal dilemma. Regarding its complexity, our data showed that the parameters responsible for the difficulty of a dilemma are, on the one hand, the ability of the participant for answering, meaning her capacity of changing her decisions depending on contextual non

moral information, and, on the other hand, the type of item, meaning how many subjects have changed their decisions depending on context in that particular dilemma (figure 36).

The classification of dilemmas between personal and impersonal is crucial for DPM approach. However, in our opinion, this distinction does not provide enough evidence to demonstrate that personal and impersonal dilemmas are underlied by different neural mechanisms. The computation that we propose can explain all type of dilemmas following the same rule.

We want to highlight a very interesting feature about our dilemmas. We have found the same distribution of behavioral outcomes using two different samples (experiment 3 was conducted in Leicester, England, and experiment 4 in Alicante) (figure 21). This could suggest that the way that people perceive PA could be more general than previously thought. An interesting question related with the former result is whether or not a participant has a typical approach to respond to the different dilemmas (Barque Duran, Pothos, Yearsley, & Hampton, 2015). Our results have shown that we can find a characteristic profile for each subject (figure 20). We think that these differences are related with the effect of context in the population (figures 15, 18 & 19). An AI close to 0 suggests that this person tend to obey deontological rules regardless of the situation, while on the contrary, values close to 1 show that the person is easily influenced by context. The distribution of our sample is centered around a mixed profile.

4.2 Dynamics of a Moral Decision

Much of the evidence in support of the DPM is related with RTs (Reaction Time). Traditionally, RTs have been used as one of the main markers for establishing the difference between automaticity and controlled processes, where the latter should take longer.

Greene originally reported an interaction between RTs and type of dilemma, such that responses affirming the moral appropriateness of harming or killing in personal dilemmas (depending on utilitarian computations) were associated with significantly longer RTs than disavowing such actions (based on deontological considerations) (Greene, et al., 2001, 2004). However, these differences were not maintained, for example, if the stimuli used were classified in a different way, more controlled and extensive (Moore, Clark, & Kane, 2008); or dilemmas which were not accepted by at least the 5% of the participants were removed (McGuire, et al., 2009; Greene, et al., 2009); or when the cognitive load in non-utilitarian and in utilitarian judgments was manipulated (Greene, et al., 2008).

Considering that these results depend on how the data was analyzed, RTs have not provided enough evidence for supporting the current DPM. In keeping with this, we have not found any significant differences in RTs between types of responses (figure 26) and this is expected since, in our hands, they are all produced by a single brain process: the estimation of Perceived Affect (figures 13 & 14). However, the analysis performed comparing the RTs in response to dilemmas classified as personal following Greene's criteria (Greene, et al., 2001), showed an interaction between the timing of acceptability and the type of response, Yes or No

(figure 34). This is an expected result, but these RTs' differences not necessarily reflect a different brain mechanism.

Our data reveals that 50% of the Deontological and the Contextual responses had latencies around 2 seconds, specifically 2.03 seconds in the case of DR and 2.16 seconds in the case of CR (figure 26B). Responses to visual control questions, obtained in experiment 3, showed similar values M= 2.211 SD= 1.706. This means that participants were extremely fast in making moral decisions (figure 30) as fast as in making perceptual decisions. This could suggest that both perceptual and moral judgments, regardless of their behavioral output, deontological or contextual, were the product of a similar underlying neural mechanism (figures 12 & 13).

The reason why this type of judgment is so fast could be because we are experts estimating affect. Affect means in this case, as we have mentioned before, the specific quality of "goodness" or "badness" experienced as a feeling state (with or without consciousness) and demarcating a positive or negative quality of a stimulus (Slovic, 2002). Affective responses occur rapidly and automatically, not only in response to moral stimuli, but also with any kind of event. The work of Kahneman and Tversky (1979) and Kahneman, Slovic, and Tversky (1982) demonstrated how individuals use many types of heuristics to make judgments (Kahneman & Tversky, 1974).

But why some trials, particularly those trials that precede the first contextual response, require more effort in terms of time and also in cognitive load? We think that in these situations, subjects are suffering a "scaffolding" process (Hogan, & Pessley, 1997; Slovic, 2000). This is a natural procedure through which new concepts are formed. Features of less usual (understood or unexpected) concepts are mapped onto existing and well-understood ones; such that the structure of the developmentally earlier, primary concept is retained in the newly constructed concept. This structure imbues the newer concept with meaning, older and newer concepts become entrained. Once the new concept is integrated, the judgment is extremely fast.

A similar phenomenon is occurring with the concept cells, neurons involved in memory functions in the hippocampus and the surrounding cortex. They provide an abstract representation of our experience, and prototype information relating to people, objects and places. Each time that an event was remembered, our brains rewrite the memory to fit the former experience with the new one (Quian Quiroga, Reddy, Kreiman, Koch, & Fried, 2005).

Turning to the moral field again, let us exemplify this concept with one of our moral dilemmas:

1. A mother lets her 11-month-old baby dies.
2. The mother also has two more children, one aged 2 years old and the other one 4. They were all on their way home in the car when a strong storm broke causing a flood.
3. The mother managed to get all of them out of the car. Immediately, the flood carried the car away and she was able to find a tree to hold on to while holding onto her three children.

4. At a given moment, due to a lack of strength and the risk that the flood would take all four of them, she had to let go of the baby.

The protagonist of the scene is a mother. All of us have our own concept of a mother with a particular value of affect (SA). This probably is very distant in respect to the affect attributed to the protagonist in trial 1, where the mother apparently makes an extremely unusual act. Because the distance, in terms of PA, between our stereotype and the affect evoked in trial 1 is large, the judgment is made rapidly. As we add more and more context, this distance, between our own stereotype (which is changing every time that we vary the situational factors) and the protagonist in the current trial, decreases. But why are RTs longer in trials where the SA and the PA are closer? This question is still open for discussion (figure 30).

As in the case of the timing explanation, choices can be flexible, if the information exposed is congruent or incongruent with the expectations, the judgment will be made very rapidly. If the information exposed is, on the contrary, incongruent with the expectations, but not distant, the participant needs more time to integrate the information into his/her internal representation.

The scaffolding phenomenon is related to one of the executive functions, the cognitive flexibility that has been described as the ability to change what you are thinking about, how you are thinking about it and even what you think about it, in other words, the ability to change your thoughts (Kaufman, 2013).

What we propose is that the different clusters (C1, C2, C3, C4 and D) evidence diverse degrees of cognitive flexibility. While the deontological participants would show less flexibility, the C1 ones would be the most flexible (figure 15). This implies a difference, in terms of time, suggesting different latencies in crossing the threshold.

Finally, the idea that we want to transmit is different from the dominant perspective in moral Psychology. Our result, instead of demonstrating that emotion and reason are opposing and parallel forces, suggests that both types of cognitive processes interact in moral decision making, depending on the contrast between SA and PA. The different RTs' between deontological or contextual responses do not necessarily require two different types of underlying neural processes. We can explain both using the same mechanism: context dependent affect gain control.

5. Final Remarks

Common human moral judgments are rife with apparent inconsistency. DPM are the current dominant class of theories of human decision making that argue for the existence of two separate, opposing decision systems. Thus DPMs posit that different situations are judged differently, because the emotional system is more strongly engaged for deontological responses while the rational system is involved in utilitarian (contextual) responses.

Our results failed to demonstrate the classical differences that relate type of response (deontological or contextual) and neural system (automatic vs rational). Low cognitive load or automatic responses mainly occurs in decisions that corroborate the acceptance of a previously evaluated moral scene, while high cognitive load or more elaborated responses, involve trials where the responses were either deontological or the first contextual.

What we are describing is not necessary incompatible with the notion of two different processes. As our results have demonstrated, both types of cognitive processes interact in moral decision making. Engaging the rational or emotional system depends on who is confronting the dilemma and her relation with the contextual information. Focusing on the person and not on the type of dilemma, let us understand which is the parameter, and therefore the mechanism, responsible of moral decision making: a subjective value representation, defined as the contrast between the Affect evoked by the protagonist of the dilemma (SA) and the influence of the situational factors (CA).

These results suggest a fuller understanding of the moral mechanism. Now we know how moral responses vary through the lens of a relatively simple, widespread and computationally precise mechanism: the divisive normalization.

UNIVERSITAS
Miguel
Hernández

Conclusions

1. Mechanism in Moral Decision Making:

    1.1 Perceived Affect is the value that underlies Moral choices. Perceived Affect as Final Common Pathway.

    1.2 Perceived Affect depends on situational context: not only on the situational factors, but also on the previous experience.

    1.3 Divisive Normalization and Affect gain control are the mechanisms responsible for moral decisions, suggesting that DN can be considered a true canonical operation of brain circuits

2. Morality at the Population Level:

    2.1 Participants have a usual way to respond to moral dilemmas suggesting different levels of contextual susceptibilities.

    2.2 Men and women perceive affect differently the PA. Men and women have different Acceptability ratings.

3. Dynamics of Moral Decisions:

    3.1 Low cognitive loads mainly occur in decisions that corroborate the acceptance of a previously evaluated moral scene, while high cognitive loads are mainly involved in those trials when decisions first departed from firm deontological rules.

    3.2 There were no statistical differences in RTs between Contextual and Deontological responses.

4. The participants' behavior could be predicted above chance from their patterns of eye movements while responding to our moral dilemmas.

5. Using signal-detection theory we have shown that subjects were rarely unware of their own deliberative process, suggesting that post-hoc mechanism do not override the conscious perception of the choice.

6. Our results did not differ between personal and impersonal dilemmas, as it would be predicted by the classical DPM (Greene, et al., 2001, 2004)

7. Participants' ability and the difficulty of the dilemmas can be used as predictors of behavioral output.

Conclusiones

1. Mecanismo Responsable de la Toma de Decisiones Morales:

    1.1 El Afecto Percibido (PA) es el valor que subyace a las decisiones morales. El PA es la vía final común.

    1.2 El PA es dependiente del contexto, no sólo de los factores situacionales, sino también de la experiencia previa.

    1.3 El control de ganancia y la normalización divisiva del afecto son los mecanismos responsables de la toma de decisiones morales.

2. Características Morales de la Población:

    2.1 Los participantes tienen una manera usual de responder a los dilemas morales, lo que sugiere que el contexto influye de manera diferente a los individuos.

    2.2 Hombres y mujeres perciben de manera diferente el PA y tienen diferentes proporciones de aceptabilidad.

3. Patrón dinámico de la Toma de Decisiones Morales:

    3.1 La baja carga cognitiva ocurre principalmente en las decisiones que corroboran la aceptabilidad de una escena moral, mientras que la alta carga cognitiva ocurre en los trials con respuesta deontológica.

    3.2 No hay diferencias significativas entre los tiempos de reacción de respuestas contextuales y deontológicas.

4. La conducta de los participantes se puede predecir por encima de un nivel de azar a partir de los movimientos oculares mientras los sujetos responden a los dilemas.

5. Utilizando la teoría de detección de señales, hemos demostrado que los sujetos normalmente son conscientes del proceso de deliberación de la decisión.

6. Nuestros resultados no difieren entre dilemas personales e impersonales como predice el DPM (Greene, et al., 2001, 2004).

7. La habilidad de los sujetos y la dificultad del dilema se pueden considerar como predictores de la conducta moral.

Bibliography

Abrams, D., Rutland, A. & Cameron, L. (2003). The development of subjective group dynamics: children's judgments of normative and deviant in-goup and out-group individuals. *Child Development, 74*(6), 1840-1856.

Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nat Rev Neurosci, 15*(10), 670-682.

Asch, S. E. (1956). Studies of independence and conformity. A minority of one against a unanimous majority. *Psychological Monographs*, 70(9), 1–70.

Bargh J. A. & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*(7), 462-479.

Bargh, j. A., Chen, M. & Burrows, L. (1996). Automaticity of social behavior: Direc effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71*(2), 230-244.

Baron, J. & Ritov, I. (2009). Psychology of Learning and Motivation vol 64 (Ed. Ross, B. H.). New York: Academic Press.

Baron, J., & Ritov, I. (2004). Omission bias, individual differences, and normality. *Organizational Behavior and Human Decision Processes, 94*(2), 74-85.

Baron, J., Gürçay, B., Moore, A. B., & Starcke, K. (2012). Use of a Rasch model to predict response times to utilitarian moral dilemmas. *Synthese, 189*(S1), 107-117.

Barque-Duran, A., Pothos, E. M., Yearsley, J. M., & Hampton, J. A. (2015). Patterns and evolution of moral behaviour: moral dynamics in everyday life. *Thinking & Reasoning, 22*(1), 31-56.

Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition, 108*(2), 381-417.

Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition, 121*(1), 154-161.

Baumard, N., & Boyer, P. (2013). Explaining moral religions. *Trends Cogn Sci, 17*(6), 272-280.

Baumard, N., Hyafil, A., Morris, I., & Boyer, P. (2015). Increased affluence explains the emergence of ascetic wisdoms and moralizing religions. *Curr Biol, 25*(1), 10-15.

Bennis, W. M., Medin, D. L. & Bartels, D. M. (2010). The costs and benefits of calculation and moral rules. Perspectives on Psychological Science, 1-16.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. (2006). The physics of optimal decision making: a formal analysis of performance in two-alternative forced-choice tasks. *Psychol Rev. 113*(4): 700-765.

Borg, J. S., Lieberman, D., & Kiehl, K. A. (2008). Infection, Incest, and Iniquity: Investigating the Neural Correlates of Disgust and Morality. *Journal of Cognitive Neuroscience*, *20*(9), 1529–1546.

Brewer, M. B. (1999). The Psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55*(39) 429-444.

Brugger, P. (1999). ONE HUNDRED YEARS OF AN AMBIGUOUS FIGURE: HAPPY BIRTHDAY, DUCK/RABBIT! *Perceptual and Motor Skills, 89*(3), 973-977.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nat Rev Neurosci, 13*(1), 51-62.

Christensen, J. F., Flexas, A., Calabrese, M., Gut, N. K., & Gomila, A. (2014). Moral judgment reloaded: a moral dilemma validation study. *Front Psychol, 5*, 607.

Christensen, J. F., & Gomila, A. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: A principled review. *Neuroscience & Biobehavioral Reviews*, 36(4), 1249-1264.

Ciaramelli, E., Muccioli, M., Ladavas, E., & di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Soc Cogn Affect Neurosci, 2*(2), 84-92.

Cikara, M., Farnsworth, R. A., Harris, L. T., & Fiske, S. T. (2010). On the wrong side of the trolley track: neural correlates of relative social valuation. *Soc Cogn Affect Neurosci, 5*(4), 404-413.

Cleckley, H. M. (1941). *The mask of sanity* (5th ed.). Georgia: Mosby Medical Library.

Colby, A. & Kohlberg, L. (1987). The measurement of moral judgment. Vol 1: Theoretical foundations and research validation. New York: Cambridge University Press.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. *Behav Res Methods Instrum Comput, 34*(4), 613-617.

Crockett, M. J. (2013). Models of morality. *Trends Cogn Sci, 17*(8), 363-366.

Crockett, M. J. (2016). How Formal Models Can Illuminate Mechanisms of Moral Judgment and Decision Making. *Current Directions in Psychological Science, 25*, 85-90.

Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences, USA, 111,* 17320-17325.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map. *40*, 61-149.

Cuddy, A. J., Fiske, S. T., Kwan, V. S., Glick, P., Demoulin, S., Leyens, J. P., & Ziegler, R. (2009). Stereotype content model across cultures: towards universal similarities and some differences. *Br J Soc Psychol, 48*(Pt 1), 1-33.

Cushman, F. A., & Greene, J. D. (2012). Finding faults: How moral dilemmas illuminate cognitive structure. *Social Neuroscience, 7*(3), 269-279.

Cushman, F. A., & Young, L. (2010) Patterns of moral judgment derive from nonmoral psychological representations. *Cognitive Science, 35*, 1052-1075.

Cushman, F. A., Young, L., & Hauser, M. D. (2006b). **T**he Role of Conscious Reasoning and Intuition in Moral Judgments: Testing three principles of harm**.** *Psychological Science 17*(12), 1082-1089.

Cushman, F., Murray, D., Gordon-McKeon, S., Wharton, S., & Greene, J. D. (2012). Judgment before principle: engagement of the frontoparietal control network in condemning harms of omission. *Soc Cogn Affect Neurosci, 7*(8), 888-895.

Cushman, F., Young, L., & Greene, J. D. (2010). *The Oxford Handbook of Moral Psychology* (eds. J. Doris, G. Harman, S. Nichols, J. Prinz, W. Sinnott-Armstrong, and S. Stich). New York: Oxford University Press.

Damasio, A. (1994). *Descarte's Error*. New York: Avon Books.

De Rivera, L. & Abuín, M. R. (2012). *Listado de Síntomas Breves (LSB-50).* Madrid: TEA Ediciones. ,

De Waal, F., (1996). *Good Natured: the origins of right and wrong in humans and other animals.* London: Harvard University Press

Derogatis, L. R. (1992). *Brief Symptom Inventory (BSY).* San Antonio: Pearson.

DeScioli, P., Bruening, R., & Kurzban, R. (2011). The omission effect in moral cognition: toward a functional explanation. *Evolution and Human Behavior, 32*(3), 204-215.

Durkheim, E. (1973). *Moral Education.* London: Collier Macmillan

Edwards, K., & von Hippel, W. (1995). Hearts and minds: The priority of affective versus cognitive factors in person perception. *Personality and Social Psychology Bulletin, 21*, 996-1011.

Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist, 49*(8), 709-724.

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: the relationship between real and hypothetical moral choices. *Cognition, 123*(3), 434-441.

Feltz, A., & Cokely, E. T. (2009). Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism. *Conscious Cogn, 18*(1), 342-350.

Fiddick, L., Spampinato, M. V., & Grafman, J. (2005). Social contracts and precautions activate different neurological systems: an fMRI investigation of deontic reasoning. *Neuroimage, 28*(4), 778-786.

Fine, C. (2006). Is the emotional dog wagging its rational tail, or chasing it? *Philosophical Explorations, (9)*1, 83-98.

Fiske S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878-902.

Fleming, S, M., & Lau, H., C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience, 8*(443), 1-9.

Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, *5*, 5-15.

Frijda, N. H. (1988). The Laws of Emotion. *American Psychologist, 43*(5), 349-358.

Gleichgerrcht, E., & Young, L. (2013). Low levels of empathic concern predict utilitarian moral judgment. *Plos One, 8*(4).

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annu Rev Neurosci, 30*, 535-574.

Graham, J., Nosek, B. A., & Haidt, J. (2011). Mapping the Moral Domain. *Journal of Personality and Social Psychology, 101* (2), 366 –385.

Greene, J. D. (2015). The rise of moral cognition. *Cognition, 135*, 39-42.

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: the interaction between personal force and intention in moral judgment. *Cognition, 111*(3), 364-371.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008).Cognitive load interferes with utilitarian moral judgment. *Cognition*, 107, 1144-1154.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44*(2), 389-400.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen J. D. (2001).An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*, 2105-2108.

Greene, J.D. (2009). Fruit flies of the moral mind. In *What's Next: Dispatches from the Future of Science*, M. Brockman, Ed., Vintage, New York.

Greene, J. D. (2007). *Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Disease, and Development* (ed. W. Sinnott-Armstrong). Cambridge: MIT Press.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-eestem and stereotypes. *Psychological Review, 102*(1), 4-27.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814-834.

Haidt, J., & Bjorklund, F. (2007). Social intuitionists answer six questions about morality. In W. Sinnott-Armstrong (Ed.), Moral psychology, Vol. 2: *The cognitive science of morality* (pp. 181-217). Cambridge, MA: MIT Press.

Haidt, J., & Hersh, M. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology,31,* 191-221.

Haidt, J., & Joseph, C. (2004). Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues. *Daedalus*, 55-66.

Haidt, J., Bjorklund, F., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no reason. *Unpublished manuscript, University of Virginia*.

Hare, R. D. (2003). *Manual for the Revised Psychopathy Checklist* (2nd ed.). Toronto, ON, Canada: Multi-Health Systems.

Hauser, M. D., Cushman, F. A., Young, L., Kang-Xing Jin, R., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind and Language 22*(1), 1-21.

Hogan, K., & Pressley, M. (1997). *Scaffolding student learning: Instructional approaches and issues*. Cambridge, MA: Brookline Books.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, 79*(8), 2554 –2558

Hume, D. A. (1978). *A treatise of human nature* (eds. Selby-Bigge, L.A. & Nidditch, P. H.). Oxford: Clarendon. (Original work published on 1739).

Iliev, R., Sachdeva, S., & Medin, D. (2012) Moral kinematics: The role of physical factors in moral judgments. *Memory and Cognition*, Online.

Immanuel, K. (1964). *Groundwork of the Metaphisic of Morals* (translated by Patton, H.J.). New York: Harper and Row. (Original work published on 1785).

Isen, A. M., & Levin, P. F. (1972). The effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology, 21*, 384-388.

Kahane, G. (2013). The armchair and the trolley. *Philosophical Studies, 162*(2), 421-445.

Kahneman, D. & Tversky, A. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(2), 1124-1131.

Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica, 47*(4157), 263-292.

Kahneman, D., & Tversky, A. (1981). The framing of decisions and the Psychology of choice. *Science, 211*, 453-458.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review, 93*(5), 1449-1475.

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

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under Uncertainty. Heuristics and Biases.* New York: Cambridge University Press

Kang, M. J., Rangel, A., Camus, M., & Camerer, C. F. (2011). Hypothetical and real choice differentially activate common valuation areas. *J Neurosci, 31*(2), 461-468.

Kaufman, J.C., Kaufman, S.B., & Plucker, J.A. (2013). Contemporary theories of intelligence. (Ed. Reisberg, J.). The Oxford Handbook of Cognitive Psychology (pp. 811-822). New York, NY: Oxford University Press.

Keltner, D., & Lerner, J. S. (2010). Emotion *Handbook of Social Psychology*. NJ: John Wiley & Sons.

Kitaoka, A. (2016). Akiyoshi's illusions pages. [online] Available at: http://www.ritsumei.ac.jp/~akitaoka/index-e.html.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature, 446*(7138), 908-911.

Kohlberg, L. (1958). *The Development of Modes of Thinking and Choices in Years 10 to 16*. Ph. D. Dissertation, University of Chicago.

Kohlberg, L. (1969). Stage and sequence: the cognitive-developmental approach to socialization. In D. A. Goslin (ed.), *Handbook of socialization theory and research* (pp. 347-480). Chicago: McNally

Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away in the study of moral development. In T. Mischel (Ed.), *Psychology and genetic epistemology* (pp. 151-235). New York: Academic Press.

Kohlberg, L. (1984). *The Psychology of Moral Development: The Nature and Validity of Moral Stages (Essays on Moral Development, Volume 2).* San Francisco: Harper & Row.

Kohlberg, L., Levine, C., & Hewer, A. (1983). *Moral stages: A current formulation and a response to critics*. Basel: Karger.

Kunda, Z. (1990). The Case for Motivated Reasoning. *American Psychologist, 108*(3), 480-498.

Kvaran, T., & Sanfey, A. G. (2010). Toward an integrated neuroscience of morality: the contribution of neuroeconomics to moral cognition. *Top Cogn Sci, 2*(3), 579-595.

Lafer-Sousa, R., Hermann, K. L., & Conway, B. R. (2015). Striking individual differences in color perception uncovered by 'the dress' photograph. *Curr Biol, 25*(13), R545-546.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annu Rev Psychol, 66*, 799-823.

Loewestein, G., & Lerner, J. S. (2003). *Handbook of affective science* (eds. Davidson, R. J., Scherer, K. R. & Goldsmith, H. H). New York: Oxford University Press.

Loewestein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feeling. *Psychological Bulletin, 127*(2), 267-286.

Louie, K., Grattan, L. E., & Glimcher, P. W. (2011). Reward value-based gain control: divisive normalization in parietal cortex. *J Neurosci, 31*(29), 10627-10639.

Louie, K., Khaw, M. W., & Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *Proc Natl Acad Sci U S A, 110*(15), 6139-6144.

Louie, K., LoFaro, T., Webb, R., & Glimcher, P. W. (2014). Dynamic divisive normalization predicts time-varying value coding in decision-related circuits. *J Neurosci, 34*(48), 16046-16057.

McCullagh, P., & Nelder, J. A. (1982). *Generalized Linear Models* (2nd ed.). London: Chapman and Hall/CRC Press.

McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models* (2nd ed.). Boca Raton, Florida: Chapman & Hall/CRC.

McGuire, J., Langdon, R., Coltheart, M., & Mackenzie, C. (2009). A reanalysis of the personal/impersonal distinction in moral psychology research. *Journal of Experimental Social Psychology, 45*(3), 577-580.

Mikhail, J. M. (2007). Universal moral grammar: theory, evidence and the future. *Trends Cogn Sci, 11*(4), 143-152.

Mikhail, J. M. (2000). *Rawls' linguistic analogy: a study of the "generative grammar" model of moral theory described by John Rawls in a Theory of Justice*. Ph.D. Dissertation, Cornell University.

Mill, J. S. (1998). *Utilitarianism* (eds. Crisp, R.). New York: Oxford University Press. (Original work published on 1861).

Moll, J., De Oliveira-Souza, R., & Zahn, R. (2008). The neural basis of moral cognition: sentiments, concepts, and values. *Ann N Y Acad Sci, 1124*, 161-180.

Moore, A. B., Clark, B. A., & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science, 19(6)*, 549-557.

Moretto, G., Làdavas, E., Mattioli, F., & di Pellegrino, G. (2009). A Psychophysiological Investigation of Moral Judgment after Ventromedial Prefrontal Damage. *Journal of Cognitive Neuroscience, 22*(8), 1888-1899.

Müller-Lyer, F. C. (1889). The Müller-Lyer illusion. [online] Available at: https://www.rit.edu/cla/gssp400/muller/muller.html.

Navajas, J., Ahmadi, M., & Quian Quiroga, R. (2013). Uncovering the mechanisms of conscious face perception: a single-trial study of the n170 responses. *J Neurosci, 33*(4), 1337-1343.

Neusner, J., & Chilton B. D. (2008). *The Golden Rule. The Ethics of Reciprocity in World Religions.* London: Continuum.

Nisbett, R. E., & Wilson, T. D. (1977). Telling More Than We Can Know: Verbal Reports on Mental Processes. *Psychological Review, 84*(3), 231-260.

Olsen, S. R., Bhandawat, V., & Wilson, R. I. (2010). Divisive normalization in olfactory population codes. *Neuron, 66*(2), 287-299.

Osgood, C. E. (1962). Studies on the generality of affective meaning systems. *American Psychologist*, *17*(1), 10-28.

Paharia, N., Kassam, K. S., Greene, J. D., & Bazerman, M. H. (2009). Dirty work, clean hands: The moral psychology of indirect agency. *Organizational Behavior and Human Decision Processes, 109*(2), 134-141.

Paxton, J. M., & Greene, J. D. (2010). Moral reasoning: hints and allegations. *Top Cogn Sci, 2*(3), 511-527.

Petrinovich, L., O'Neill, P., & Jorgensen, M. J. (1993). An empirical study of moral intuitions: Towards an evolutionary ethics. *Journal of Personality and Social Psychology*, *64*(3), 467–478.

Phelps, E. A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annu Rev Psychol, 57*, 27-53.

Phelps, E. A., Lempert, K. M., & Sokol-Hessner, P. (2014). Emotion and decision making: multiple modulatory neural circuits. *Annu Rev Neurosci, 37*, 263-287. Quian Quiroga, R., & Panzeri, S. (2009). Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci, 10*(3), 173-185.

Piaget, J. (1932). *The moral judgment of the child*. London: Kegan Paul, Trench, Trubner & Co.

Proffitt, D. R. (2006). Distance perception. *Current Directions in Psychological Research, 15*, 131-135.

Quian Quiroga, R., & Panzeri, S. (2009). Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci, 10*(3), 173-185.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature, 435*(7045), 1102-1107.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci, 9*(7), 545-556.

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Comput. 20*(4), 873–922.

Rest, J. R. (1980). Development in moral judgment research. *Developmental Psychology, 16*(4), 251-256.

Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron, 61*(2), 168-185.

Royzman, E. B., & Baron, J. (2002). The preference forindirect harm. *Social Justice Research, 15,* (2), 165-184.

Saltzstein, H. D., & Kasachkoff, T. (2004). Haidt's Moral Intuitionist Theory: A Psychological and Philosophical Critique. *Review of General Psychology, 8*(4), 273-282.

Saver, J. L., & Damasio, A. R. (1991).Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia*, *29*(12): 1241–1249.

Schaich Borg, J., Hynes, C., Van Horn, J., Grafton, S., & Sinnot-Armstrong, W. (2008). Consequences, action and intention as factors in moral judgments: an FMRI investigation. *Journal of Cognitive Neuroscience, 18*, 803-817.

Selman, R. L. (1971) The relation of role-taking ability to the development of moral judgment in children. *Child Development, 42*, 79-91.

Shavitt, S. (1990). The role of attitude objects in attitude functions. *Journal of Experimental social Psychology, 26*, 124-148.

Shenhav, A., & Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron, 67*(4), 667-677.

Shenhav, A., & Greene, J. D. (2014). Integrative moral judgment: dissociating the roles of the amygdala and ventromedial prefrontal cortex. *J Neurosci, 34*(13), 4741-4749.

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review, 74*(1), 29-39.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*(1), 3-22.

Slovic, P. (2000). *Perception of risk*. London: Earthscan.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). *Heuristics and biases: The Psychology of intuitive judgment* (eds. Gilovivh, T., Griffin, D. & Kahneman, D.). New York: Cambridge University Press.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D.G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis, 24*(2), 311–322.

Smith, P. L., & Vickers, D. (1988). *The Oxford Handbook of Computational and Mathematical Psychology* (eds. Busemeyer, J. R., Wang, Z., Townsend, J. T., Eidels, A.). New York: Oxford University.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*, 645-726.

Statman, D. (1997). *Virtue Ethics*. Cambridge: Edinburgh University Press.

Suter, R. S., & Hertwig, R. (2011). Time and moral judgment. *Cognition, 119*(3), 454-458.

Swann, W. B., Gómez, A., Buhrmester, M. D., & López-Rodríguez, L. (2014). Contemplating the ultimate sacrifice: identify fusion channels por-group affect, cognition and moral decision. Journal of Personality and Social Psychology, 106(5), 713-727.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science, 12,* 257-285.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behaviour. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 7-24). Chicago: Nelson-Hall.

Takeuchi, T., & De Valois, K. K. (2000). Modulation of perceived contradt by a moving surround. *Vision Research, 40*, 2649-2709.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53-55.

Teper, R., Inzlicht, M., & Page-Gould, E. (2011). Are we more moral than we think? Exploring the role of affect in moral behavior and moral forecasting. *Psychol Sci, 22*(4), 553-558.

Thomson, J. J. (1985). The Trolley Problem. *The Yale Law Journal, 94*(6), 1395-1415.

Tremoliere, B., Neys, W. D., & Bonnefon, J. F. (2012). Mortality salience and morality: thinking about death makes people less utilitarian. *Cognition, 124*(3), 379-384.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59*, 433-460.

Tversky, A., & Simonson, I. (1993). Context-Dependent Preferences. *Management science 39* (10), 1179-1189.

Uhlmann, E. L., Pizarro, D. A., Tannenbaum, D., & Ditto, P. H. (2009). The motivated use of moral principles. *Judgment and Decision Making, 4*, 476-491.

Usher, M., & McClelland, J. L. (2001). On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review*, *108*, 550-592.

Verdonck S., Tuerlinckx F. 2016. Factoring out nondecision time in choice reaction time data: Theory and implications. *Psychological Review, 123* (2), 208-218.

Vitale, J. E., Newman, J. P., Bates, J. E., Goodnight, J., Dodge, K. A., & Pettit, G. S. (2005). Deficient Behavioral Inhibition and Anomalous Selective Attention in a Community Sample of Adolescents with Psychopathic Traits and Low-Anxiety Traits. *Journal of Abnormal Child Psychology, 33*(4), 461-470.

Wabersich, D., & Vanderkerckhove, J. (2014). The RWiener Package: an R Package Providing Distribution Functions for the Wiener Diffusion Model. *The R Journal, 6(1),* 49-56.

Waldman, M. R., & Dieterich, J. H. (2007). Throwing a bomb on a person versus thorwing a person on a bomb. Psychological Science, 18(3), 247-253.

Wells, G. I., & Windschitl, P. D. (1999). Stimulus sampling and social Psychological experimentation. *Society for Personality and Social Psychology, 25*(9), 1115-1125.

Wikipedia. (2016). The dress (viral phenomenon). [online] Available at: https://en.wikipedia.org/wiki/The_dress_(viral_phenomenon).

Wilson, E. O. (1975). *Sociobiology the New Sinthesis*. Cambridge: Harvard University Press.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A Model of dual attitudes. *Psychological Review, 107*(1), 101-126.

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin*, *24*, 1245-1257.

Zajonc, R. B. (1980). Feeling and thinking. Preferences need no inferences. *American Psychologist, 35(*2), 151-175.

Appendixes

I. Lawrence Kohlbergh's Moral Stages
II. Evidence in favour of the DPM current
III. The Classical Trolley & Footbridge Dilemmas
IV. First Affect Evidence
V. CA Definition
VI. Cronbach's Alpha Analysis
VII. Psychological Questionnaires
VIII. χ^2 Analysis: Relation between Acc and PA
IX. Analysis of the Threshold Question
X. Rasch Model
XI. Results Special Moral Dilemmas
XII. Drift diffusion Model
XIII. Individual ROC Space
XIV. Cross Validation

Appendix I – Lawrence Kohlbergh's Moral Stages

Based on his study of children's moral responses, Kohlberg expanded Piaget's three stages into six, organized into three levels, each level consisting of two stages (Kohlberg, 1958, 1963, 1969, 1971, 1984).

Level I: Preconventional Morality. Children think of morality in terms of the consequences of disobedience to adult rules in order to avoid punishment. Behaviors are "good" or "bad" depending on their consequences, or in other words, behavior is guided by rewards and punishments. The children at this stage do not comprehend the rules of society.

- Stage 1. Obedience and Punishment Orientation. The child/individual's behavior is good in order to avoid being punished. If a person is punished they must have done wrong.
- Stage 2. Individualism and Exchange. At this stage children recognize that there is not just one right view that is handed down by the authorities. Different individuals have different viewpoints.

Level II: Conventional Morality. At this level the child begins to grasp social rules and gains a more objective perspective on right and wrong.

- Stage 3. Good Interpersonal Relationships. The child/individual's behavior is good in order to be seen as being a good person by others. Therefore, answers are related to the approval of others.
- Stage 4. Maintaining the Social Order. The child/individual becomes aware of the wider rules of society so judgments concern obeying rules in order to uphold the law and to avoid guilt.

Level III: Postconventional Morality. At this level the emphasis is no longer on conventional, societal standards of morality, but rather on personal or idealized principles.

- Stage 5. Social Contract and Individual Rights. The child/individual becomes aware that while rules/laws might exist for the good of the greatest number of people, there are times when they will work against the interest of particular individuals.
- Stage 6. Universal Principles. People at this stage have developed their own set of moral guidelines which may or may not fit the law. The principles apply to everyone. E.g. human rights, justice and equality. The person will be prepared to act to defend these principles even if it means going against the rest of society in the process and having to pay the consequences of disapproval and or imprisonment. Kohlberg doubted few people reached this stage.

Appendix II – Evidence in favour of the Current DPMs

A strong support in favor of DPM came from Antonio Damasio's *Descartes' Error* (1994), a book describing studies of decision making in patients with lesions to the ventromedial prefrontal cortex (VMPFC), one of the regions damaged in the famous case of Phineas Gage. VMPFC patients were rare because their real-life decision-making ability was clearly impaired by their lesions, but their deficits typically evaded detection using neurological measures of executive function (Saver, & Damasio, 1991). Notably, such patients showed no signs of impairment on Kohlberg's widely used test of moral reasoning (Colby, & Kohlberg, 1987). It seems that VMPFC patients make poor decisions because they are unable to generate feelings that guide adaptive decision-making in healthy individuals.

Damasio showed that morality could be studied using the new technology of fMRI, and also that morality, and rationality itself, were crucially dependent on the proper functioning of emotional circuits in the prefrontal cortex.

Also infavor of DPM, Frans de Waal's *Good Natured*, a book published just two years after Damasio's, proposes that the building blocks of human morality are found in other apes and are the product of natural selection in the highly social primate lineage (de Waal, 1996).

De Waal argues that animals acquired a natural sense of morality millions of years before humans even appeared on the evolutionary scene. He wrote: "*Humans and other animals have been endowed with a capacity for genuine love, sympathy and care*." This fact "*can and will one day be reconciled with the idea that genetic self-promotion drives the evolutionary process*." (de Waal, 1996).

These two books came out just as John Bargh was showing to social psychologists that automatic and unconscious processes can and probably do cause the majority of our behaviors, even those with moral connotations (Bargh, 1996, 1999).

These findings suggested that emotion played a bigger role in moral decision making than previous rational developmental psychologists thought.

Appendix III – The Classical Trolley & Footbridge Dilemmas

Why almost everybody responds NO to the Footbridge problem and YES to the Trolley dilemma? Some have argued that the issue is whether an action directly (as in Footbridge) or indirectly (as in Trolley) causes harm (Royzman, & Baron, 2002); whether the causal focus is directed on to the trolley or the people on the track (Waldmann, & Dieterich, 2007; Iliev, Sachdeva, & Medin, 2012); whether the action is interpreted as violating a rule in the social contract (Fiddick, Spampinato, & Grafman, 2005); or whether the outcomes are viewed as gains or losses (Petrinovich, O'Neill, & Jorgensen, 1993). Under the most popular interpretation, the Footbridge dilemma elicits an automatic response based on an affective mechanism, while the Trolley problem evokes a controlled one based on a more rational mechanism (Greene, et al., 2001).

However, the use of these dilemmas is highly controversial. Still people might not accept at face value the closed-world assumptions of these scenarios, they might doubt whether a fat man's body can stop a trolley car or whether this is the only available solution (Bennis, Medin, & Bartels, 2010). Moreover, even if they accept the constraints of the problem, many people find some problems to be amusing rather than sobering, and some evidence suggests that trolley problems do not always engage the same psychological processes as other moral situations. Also, importantly, relying on any one set of stimuli is problematic; if most studies on a given topic use highly similar stimuli, researchers cannot determine whether or how common features of the stimuli influence the results (Wells, & Windschitl, 1999).

At present, there is a current trend in favor of investigating moral decisions using dilemmas with action-relevant environments where the stakes are immediate (Kang, Rangel, Camus, & Camerer, 2011), emotionally charged (Teper, Inzlicht, & Page-Gould, 2011) and tangible (FeldmanHall, Mobbs, Evans, Hiscox, Navrady, & Dalgleish, 2012). Indeed, research on the psychology of choice has shown that decisions are influenced by the environments in which they are made (Isen, & Levin, 1972) and most likely also by previous experience (Crockett, 2013).

Appendix IV – First Affect Evidence

One of the first proponents of the weight of affect in decision making was Robert Zajonc who argued that affective reactions to stimuli are often the very first reactions, occurring automatically and subsequently guiding information processing and judgment. The fundamental importance of affect has been demonstrated with a simple but striking experiment, particularly, when objects were presented to an individual repeatedly, the "mere exposure" was enough for creating a positive attitude or preference for this object (Zajonc, 1980).

According to Zajonc, all perceptions contain some affect. "*We do not just see 'a house': We see a handsome house, an ugly house, or a pretentious house*" He later adds*, "We sometimes delude ourselves that we proceed in a rational manner and weigh all the pros and cons of the various alternatives. But this is probably seldom the actual case. Quite often 'I decided in favor of X' is no more than 'I liked X' . . . We buy the cars we 'like,' choose the jobs and houses we find 'attractive,' and then justify these choices by various reasons . . ."* (Zajonc, 1980).

Zajonc introduced and demonstrated that immediate emotional response, or affect, is a basic reaction and is unavoidable and everyone has emotional reactions to events and stimuli. There are decisions that people make that benefit from more cognitive and less emotional influence, but it can be difficult to extract emotional feelings from decision making precisely because these feelings are unavoidable and difficult to articulate.

Appendix V- CA Definition

The recent viral phenomenon of #TheDress (Wikipedia, 2016) is a perfect example to define how Context influences general perception, and particularly to illustrate our definition of Contextual Affect (CA) (Figure 43). This is a peculiar photography where pixels by themselves are brown and blue but popular accounts indicates that people tend to see the dress as white/gold or blue/black. What then causes this striking individual differences?

Figure 43. Dress Picture.

The influence of two types of Contextual information are key to understand this curious visual illusion:

1st The available perceptual data. The surrounding in which the object we are looking at appear in is influencing our perception of its color. The color and other attributes of an object are never judged in isolation. Rather, they are compared with information coming from regions of the visual scene immediately surrounding that object.

2nd The internal model of the world. Color appearance depends on the spectral power of the illuminant. By default, our brains assume that the illuminant is the sun, but the sunlight changes its spectral distribution depending on time of the day. Recently, it has been shown that even though we can distinguish between midday and sunset and sunrise light, we can hardly discriminate morning and afternoon illuminations. Thus, most people show a strong internal bias for a particular daylight, influenced in part by immediate past experiences. This bias represents the brain's internal model of the illuminant.

A color percept depends on both the given available sensory information and the internal model of the world. Thus, people who expect a cool illuminant, will discount short wavelengths, and perceive the dress as white/gold; while others who favor a warm illuminant, will discount longer wavelengths, and see it as blue/black (Lafer-Soussa, Hermann, & Conway, 2015). This is an example of top-down processing, where what we is in part determined by what our brain expects.

The concept of CA we introduced is also affected by this duality:

1st The available situational factors. The context in which the action takes place. For example, a moral scene could represent a protagonist performing an action in a context of war or, on the peaceful retirement of a small village.

2nd The internal moral model of the world. The group of past experiences that influence how people perceive the action in the context of the current situational factors. For example, a person lives in the midst of a civil war, will most likely have different perceptions in comparison to a person who had never gone through such an ordeal.

Appendix VI- Cronbach's Alpha Analysis

Since our experimental design is very particular (each dilemma is composed by four different trials where protagonist, action and consequences are kept constant, while different contextual information is added in sequence) we want to test to what extent these set of sequential items were measuring a single, one-dimensional construct, such as the Perceived Affect.

Cronbach's Alpha is the most common measure of internal consistency when you have multiple Likert items in a questionnaire that form a scale and you wish to determine if the scale is reliable.

The Alpha coefficient for the four items is 0.823, suggesting that the items have relatively high internal consistency (note that a reliability coefficient of 0.70 or higher is considered "acceptable" in most cases of social science reserch (Tavakol, & Dennick, 2011).

Table (24) shows the descriptive data of each trial:

Table (24)
Cronbach's Alpha Analysis: PA Descriptive Data of trials of experiment 4

	M	SEM	n
PA Tr1	0.1501	0.009	643
PA Tr2	0.2792	0.011	643
PA Tr3	0.4059	0.013	643
PA Tr4	0.4794	0.014	643

Table (25) shows two important measures of this basic reliability analysis. First of all the "corrected item-total correlation" are the correlations between each item and the total score from the questionnaire. In a reliable scale all items should correlate with the total as in our case. The values in the column labelled "Alpha if Item is Deleted" are the values of the overall Alpha if that item isn't include in the calculation. None of the items here would substantially affect reliability if they were delated.

Table (25)
Cronbach's Alpha Analysis: Item-Total Statistics

	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
PA Tr1	0.551	0.822
PA Tr2	0.690	0.758
PA Tr3	0.732	0.734
PA Tr4	0.659	0.778

Appendix VII- Psychological Questionnaires

Psychopathy is a developmental disorder that is characterized by high levels of antisocial behavior, as well as emotional impairments such as callousness and a distorted moral behavior (Cleckley, 1941; Hare, 2003).

An environment characterized by the breakdown of social norms and regulations, disorganization, undesirable peer models, and a climate of alienation from and hostility toward the broader society appears to produce a psychopathic personality. Participants of Experiment 3 were analyzed using the following questionnaires due to the relation found between the perception of hostility and the psychopathy (Vitale, Newman, Bates, Goodnight, Dodge, & Pettit, 2005).

1. Listado de Síntomas Breves (LSB-50)

1.1 Brief description

The LSB-50 (de Rivera, & Abuín, 2012) is a 50-item self-report symptom inventory designed to reflect the psychological symptom patterns of psychiatric and medical patients as well as community no patient respondents. Each item is rated on a 5-point scale of distress (0-4) ranging from "not at all" (0) at one pole to "extremely" (4) at the other. The LSB-50 is scored and profiled in terms of nine primary symptoms and three global index of distress and one global psychopathic risk index. Even the questionnaire includes two sincerity indexes.

1.2 The Primary Dimensions

- Psicorreactivity (PR): Evaluate the sensitivity in perception of self in relation to others and in relation to the image itself, as well as ways of acting and thinking with excessive self-observation.
- Hipersensibility (Hp): Explore both, intra-interpersonal, sensitivity as the excessive focusing.
- Obssessive-Compulsive (Ob): evaluates the presence of obsessions and continuing doubts that flood the mind as well as the existence of rituals and compulsions.
- Anxiety (An): Explore the demonstrations both generalized anxiety disorder and panic boxes and phobic anxiety.
- Hostility (Hs): Evaluate reactions to loss of emotional control or continuous sudden manifestations of aggression and anger.
- Somatization (Sm): Explore the presence of somatic symptoms or bodily discomfort due to psychological processes of somatization.

- Depression (De): Evaluate the presence of characteristic depression symptoms such as sadness, hopelessness, anhedonia, energy, helplessness and guilt.

- Alteration of sleep (Su): Explore specify the presence of sleep disturbances that are relevant from the point of view of health and wellness.

- Alteration of sleep II (Su-a): Explore specify the presence of sleep disorders with manifestations of depression and anxiety scales.

1.3 Global Index

- Global Severity Index: Indicates the degree of affectation of overall psychopathology assessed.
- Number of symptoms: Indicates the number of symptoms that arise in the evaluation.
- Intensity Index: It is an index of the intensity or severity of the symptoms that the subject claims to have.
- Psychopathological Risk Index (IRPpsi): Evaluate the presence of clinical symptoms associated with devaluation, incomprehension, fear, hostility and somatization with suicidal population.

1.4 Sincerity Indexes

- Minimization (Min): It indicates whether the evaluated may be minimizing the frequency or intensity of the symptoms being treated. It may reflect a conscious or unconscious attempt to give a falsely positive image of him/her.
- Maximization (Mag): It indicates whether the evaluated may be magnifying its symptomatology, scoring with high intensity rare symptoms.

1.5 Technical Aspects

- Application Norms: Under normal circumstances, the LSB-50 requires 5 to 10 minutes to complete. Typical time for administrative instructions was 2 to 5 minutes. The question that respondents should answer was: How much that problems has distressed or bothered you during the past 7 days including today? The task was to black the circle for only one number for each problem and did not skip any items.
- Correction: the raw score obtained by the average of all the items whose composed a scale, could become in Percentile (Pc) as a result of the comparison between the accrual respondent with a large sample.
- Interpretation: the operational rule for caseness provided above states that if the respondent has a IRPsi score greater than or equal to a Pc 97 score, or if any two primary dimensions to a Pc of a 97, then the individual considered a positive diagnosis or a case.

1.6 Participants' Data

Table (26) shows the Pc obtained in the primary dimensions and in the Psychopathological Risk Index. Hs and IRPpsi scales were studied in detail to localize participants with Pc superior to 97. Any of the subjects exceeded this Pc.

Table (26)
LSB Participants' Scores

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
Min	40	95	85	25	85	65	40	90	65	85	50	20	85	90	10	50	85
Mag	20	10	5	20	10	5	10	10	20	5	20	45	10	10	10	25	20
Pr	45	25	25	60	40	60	70	45	40	50	75	95	60	50	95	65	60
Hp	30	30	25	30	40	40	60	45	40	40	70	98	45	50	45	65	45
Ob	70	25	70	80	65	70	80	60	45	70	85	80	70	80	98	70	70
An	80	50	45	30	65	70	80	30	45	70	50	85	30	80	30	70	70
Hs	90	30	20	60	30	30	30	30	55	20	30	85	60	30	60	80	60
Sm	70	35	30	80	30	30	75	20	70	30	20	95	20	30	70	55	50
De	35	10	35	75	15	35	70	10	50	25	45	45	45	15	70	60	35
Su	85	30	20	96	20	70	30	30	20	20	50	50	30	20	90	35	50
Su-a	75	15	15	90	15	55	30	15	15	10	30	45	30	15	85	55	30
IRPpsi	50	50	30	50	10	10	50	35	75	10	50	80	35	45	35	85	80

Note: Participant's codification had been exhibited in an arbitrary way.

2. Brief Symptom Inventory (BSI)

2.1 Brief Description

The BSI (Derogatis, 1993) is a 53-item self-report symptom inventory designed to reflect the psychological symptom patterns of psychiatric and medical patients as well as community no patient respondents. Each item of the BSI is rated on a 5-point scale of distress (0-4) ranging from "not at all" (0) at one pole to "extremely" (4) at the other. The BSI is scored and profiled in terms of nine primary symptoms and three global indices of distress. The three global indices, nine dimensions, and 53 items reflect the three principal levels of interpretation of the BSI, descending from general superordinate measures of psychological status, through syndrome representations, to individual symptoms.

2.2 The Primary Dimensions

- Somatization (Som): this dimension reflects distress arising from perceptions of bodily dysfunction. Items focus on cardiovascular, gastrointestinal and respiratory complaints; other systems with strong automatic mediation are included as well. Pain and discomfort of the gross musculature and additional somatic equivalents of anxiety are also components of somatization.

- Obsessive-Compulsive (O-C): includes symptoms that are often identified with the standard clinical syndrome of the same name. This measure focuses on thoughts, impulses and actions that are experienced as unremitting and irresistible by the individual, but are of an ego-alien or unwanted nature. Behavior and experienced of a more general cognitive performance deficit are also included in this measure.

- Interpersonal Sensitivity (I-S): this dimension centers on feelings of personal inadequacy and inferiority, particularly in comparison with others. Self-deprecation, self-doubt, and marked discomfort during interpersonal interaction are characteristic manifestation of this syndrome.

- Depression (Dep): the symptoms of the Depression dimension reflect a representative range of the indications of clinical depression. Symptoms of dysphoric mood and affect are represented as are lack of motivation and loss of interest in life.

- Anxiety (Anx): General signs such as nervousness and tension are included in the Anxiety dimension, as are panic attacks and feelings of apprehension and some somatic correlates of anxiety are also included as dimensional components.

- Hostility (Hos): this dimension includes thoughts, feelings, or actions that are characteristics of the negative affect state of anger.

- Phobic Anxiety (Phob): Phobic Anxiety is defined as a persistent fear response- to a specific person, place, object or situation- that is irrational and disproportionate to the stimulus and leads to avoidance or escape behavior.

- Paranoid Ideation (Par): represents paranoid behavior fundamentally as a disordered mode of thinking. The cardinal characteristics of projective thought, hostility, suspiciousness, grandiosity, centrality, fear of loss of autonomy and delusions are viewed as primary aspects of this disorder. Item selection was oriented toward representing this conceptualization.

- Psychoticism (Psy): This scale was developed to represent the construct as a continuous dimension of human experience. Items indicative of a withdrawn, isolated, schizoid lifestyle were included, as were first-rank symptoms of schizophrenia, such as thought control. This scale provides for a graduated continuum from mild interpersonal alienation to dramatic psychosis.

2.3 Global index

Three global index have been developed and added to provide more flexibility in overall assessment of the patient's psychopathological status and to provide psychometric appraisal at a third, more general level of psychological well-being

- Global Severity Index: to calculate the GSI, the sums for the nine symptom dimensions and divided by the total number of responses.

- Positive Symptom Total: the PST is derived by counting the number of items endorsed with a positive (nonzero) response.

- Positive Symptom Distress Index: the PSDI is calculated by dividing the sum of the items value by the PST.

2.4 Technical Aspects

- Application Norms: Under normal circumstances, the BSI requires 8 to 10 minutes to complete. Typical time for administrative instructions was 2 to 5 minutes. The question that respondents should answer was: How much that problems has distressed or bothered you during the past 7 days including today? The task was to black the circle for only one number for each problem and did not skip any items.

- Correction: the raw scores for the nine symptom dimensions and three global indices are converted to standardized T scores available for males and females.

- Interpretation: the operational rule for caseness provided above states that if the respondent has a GSI score greater than or equal to a T score of 63, or if any two primary dimensions to a T score of a 63, then the individual considered a positive diagnosis or a case.

2.5 Participants' Data

Table (29) shows the T scores obtained in the primary dimensions. Hos scale was studied in detail to localize participants with data superior to 63. As was mentioned in the Technical Aspects, just a paragraph above, any of the subjects was more than 63 is these two specific dimensions.

Table (27)
BSI Participants' Scores

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
SOM	49	54	0	49	49	56	0	62	49	60	52	0	52	0	61	65	49	61
O-C	58	69	67	69	56	56	64	58	56	64	50	61	58	53	71	clinical	63	74
I-S	55	54	59	68	59	50	55	59	0	59	50	0	50	0	62	clinical	59	54
DEP	49	60	69	70	49	49	54	57	49	60	0	0	0	0	67	clinical	60	73
ANX	45	48	0	69	52	52	52	0	0	59	0	0	0	0	59	clinical	58	65
HOS	52	50	62	62	56	64	53	50	0	50	0	0	60	0	62	50	60	59
PHOB	0	0	0	61	0	61	0	0	0	0	0	0	0	0	0	clinical	64	60
PAR	49	47	49	49	43	43	49	0	49	42	0	0	0	0	0	47	49	47
PSY	49	62	70	70	49	55	49	0	56	69	0	0	56	69	62	69	65	69

Note: Participant's codification had been exhibited in an arbitrary way.

Appendix VIII- χ^2 Analysis: Relation between Acc and PA

Tables 28, 29, 30 & 31 show the relationship between Acceptability and Perceived Affect for the sample of experiment 4. We have the outcome of Acc tabulated by type level of PA. This kind of cross-tabulation of frequencies is also called a contingency table or cross classification. Each entry in the table is a percentage.

Table (28)
Contingence table trial 1: Acc & PA

PA_Tr1 * Accep _Tr1 Crosstabulation

			Accep Tr1		Total
			NO	YES	
PA_Tr1	-1	Count	414	1	415
		% PA	99.8%	0.2%	100.0%
	-0.5	Count	109	10	119
		% PA	91.6%	8.4%	100.0%
	0	Count	56	15	71
		% PA	78.9%	21.1%	100.0%
	0.5	Count	12	15	27
		% PA	44.4%	55.6%	100.0%
	1	Count	5	6	11
		% PA	45.5%	54.5%	100.0%
Total		Count	596	47	643
		% PA	92.7%	7.3%	100.0%

From these results, we can observe that there is relationship between PA and Acc in tr1, $\chi^2 = 179.808$; df = 4; $p < 0.0001$ (figure 44).

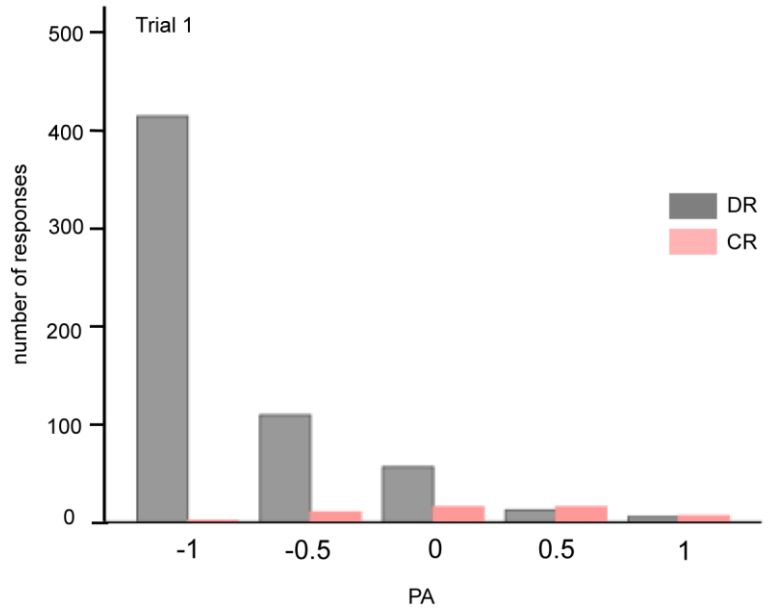


Figure 44. Relation between PA and Acc in trial 1. Bars Graph showing the number of both type of responses (Acc or no Acc) per PA level.

Table (29)
Contingence table trial 2: Acc & PA

PA_Tr2 * Accep _Tr2 Crosstabulation					
			Accep Tr2		Total
			NO	YES	
PA_Tr2	-1	Count	270	1	271
		% PA	99.6%	0.4%	100.0%
	-0.5	Count	141	17	158
		% PA	89.2%	10.8%	100.0%
	0	Count	79	39	118
		% PA	66.9%	33.1%	100.0%
	0.5	Count	19	41	60
		% PA	31.7%	68.3%	100.0%
	1	Count	4	32	36
		% PA	11.1%	88.9%	100.0%
Total		Count	513	130	643
		% PA	79.8%	20.2%	100%

From these results, we can say that there is relationship between PA and Acc in tr2, χ^2 = 278.365; df = 4; p < 0.0001 (figure 45).

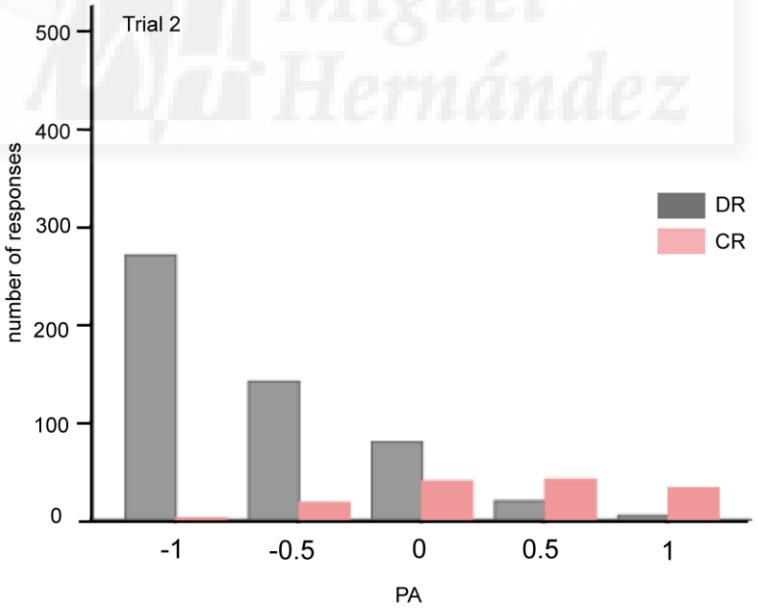


Figure 45. Relation between PA and Acc in trial 2. Bars Graph showing the number of both type of responses (Acc or no Acc) per PA level.

Table (30)
Contingence table trial 3: Acc & PA

PA_Tr3 * Accep _Tr3 Crosstabulation					
			Accep Tr3		Total
			NO	YES	
PA_Tr3	-1	Count	199	5	204
		% PA	97.5%	2.5%	100%
	-0.5	Count	96	14	110
		% PA	87.3%	12.7%	100%
	0	Count	72	55	127
		% PA	56.7%	43.3%	100%
	0.5	Count	44	84	128
		% PA	34.4%	65.6%	100%
	1	Count	10	64	74
		% PA	13.5%	86.5%	100%
Total		Count	421	222	643
		% PA	65.5%	34.5%	100%

From these results, we can observe that there is relationship between PA and Acc in tr3, $\chi^2 = 263.444$; df = 4; $p < 0.0001$ (figure 46).

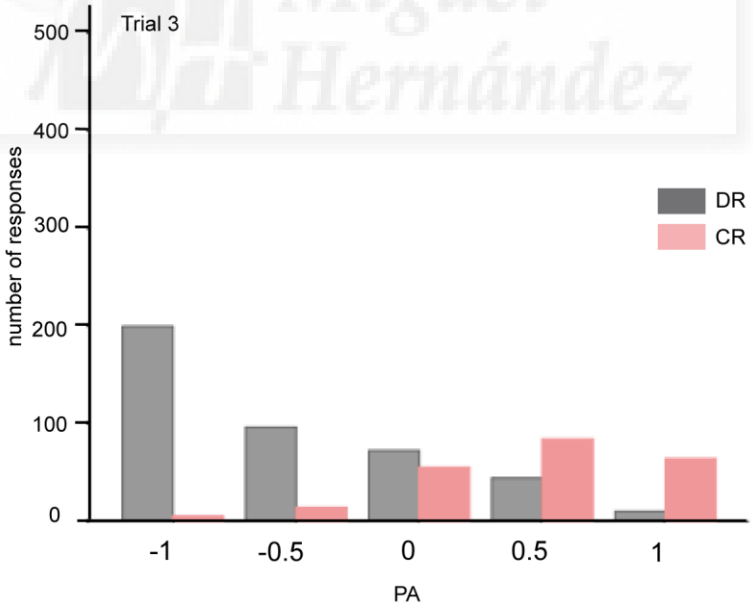


Figure 46. Relation between PA and Acc in trial 3. Bars Graph showing the number of both type of responses (Acc or no Acc) per PA level.

Table (31)
Contingence table trial 4: Acc & PA

PA_Tr4 * Accep _Tr4 Crosstabulation					
			Accep Tr4		Total
			NO	YES	
PA_Tr4	-1	Count	173	10	183
		% PA	94.5%	5.5%	100.0%
	-0.5	Count	57	19	76
		% PA	75%	25%	100.0%
	0	Count	64	56	120
		% PA	53.3%	46.7%	100.0%
	0.5	Count	38	101	139
		% PA	27.3%	72.7%	100.0%
	1	Count	12	113	125
		% PA	9.6%	90.4%	100.0%
Total		Count	344	299	643
		% PA	53.5%	46.5%	100%

From these results, we can determine that there is relationship between PA and Acc in tr4, $\chi^2 = 273.069$; df = 4; $p < 0.0001$ (figure 47).

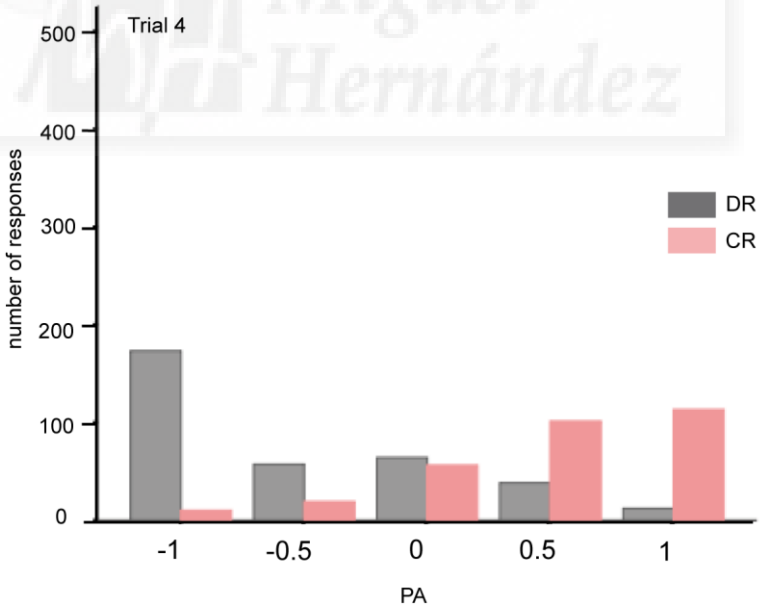


Figure 47. Relation between PA and Acc in trial 4. Bars Graph showing the number of both type of responses (Acc or no Acc) per PA level.

Appendix IX- Analysis of the Threshold Question

How do PA scores differ between the different Clusters in each trial? We have conducted a Kruskal-Wallis for each trial to address this question.

- Trial 1 General Kruskal-Wallis test:
  - N= 643
  - K= 174.937
  - df= 4
  - Pvalue= <0.0001

Comparison between Clusters is shown in table (32).

Table (32)
Clusters' Comparison tr 1

Cluster's Comparison	K	Pvalue
D vs C4	26.461	> 0.05
D vs C3	-56.791	>0.05
D vs C2	-125.227	<0.0001
D vs C1	-286.457	<0.0001
C4 vs C3	-78.253	<0.05
C4 vs C2	-151.688	<0.0001
C4 vs C1	-312.918	<0.0001
C3 vs C2	-73.435	<0.05
C3 vs C1	-234.665	<0.0001
C2 vs C1	-161.230	<0.0001

- Trial 2 General Kruskal-Wallis test:
  - N= 643
  - K= 256.342
  - df= 4
  - Pvalue= <0.0001

Comparison between Clusters is shown in table (33).

Table (33)
Clusters' Comparison tr 2

Cluster's Comparison	K	Pvalue
D vs C4	-19.988	> 0.05
D vs C3	-97.728	<0.0001
D vs C2	-274.231	<0.0001
D vs C1	-299.502	<0.0001
C4 vs C3	-77.740	<0.05
C4 vs C2	-254.243	<0.0001
C4 vs C1	-279.514	<0.0001
C3 vs C2	-176.503	<0.0001
C3 vs C1	-201.774	<0.0001
C2 vs C1	-25.271	> 0.05

- Trial 3 General Kruskal-Wallis test:
  - N= 643
  - K= 260.938
  - df= 4
  - Pvalue= <0.0001

Comparison between Clusters is shown in table (34).

Table (34)
Clusters' Comparison tr 3

Cluster's Comparison	K	Pvalue
D vs C4	-17.186	> 0.05
D vs C3	-209.529	<0.0001
D vs C2	-258.109	<0.0001
D vs C1	-274.604	<0.0001
C4 vs C3	-192.343	<0.0001
C4 vs C2	-240.923	<0.0001
C4 vs C1	-257.428	<0.0001
C3 vs C2	-48.580	>0.05
C3 vs C1	-65.075	>0.05
C2 vs C1	-16.495	>0.05

- Trial 4 General Kruskal-Wallis test:
  - N= 643
  - K= 283.268
  - df= 4
  - Pvalue= <0.0001

Comparison between Clusters is shown in table (35).

Table (35)
Clusters' Comparison tr 4

Cluster's Comparison	K	Pvalue
D vs C4	26.461	<0.0001
D vs C3	-56.791	<0.0001
D vs C2	-125.227	<0.0001
D vs C1	-286.457	<0.0001
C4 vs C3	-78.253	>0.05
C4 vs C2	-151.688	<0.05
C4 vs C1	-312.918	>0.05
C3 vs C2	-73.435	>0.05
C3 vs C1	-234.665	>0.05
C2 vs C1	-161.230	>0.05

Appendix X- Rasch Model

The following tables 36, 37 & 38 show the infit values for the three different Rash Models.

Table (36)
Infit Values for 1-10 dilemmas of experiment 3

Dilemma	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Infit	0.91	0.88	0.94	0.87	0.76	0.86	0.86	0.91	0.85	0.92

Table (37)
Infit Values for trials of dilemma 1

Dilemma 1	**1**	**2**	**3**	**4**
Infit	0.47	0.53	0.73	0.76

Table (38)
Infit Values for trials of dilemma 7

Dilemma 7	**1**	**2**	**3**	**4**
Infit	0.47	0.77	0.97	1

Appendix XI- Special Dilemmas

We have introduced 5 special dilemmas where the PA manipulations were in both directions (increasing and decreasing) introducing more than one change in the participants' acceptability judgments. Moreover, Scenes 11 and 13 had the Deontological Response in the different direction, meaning that participants who gave a DR answered clicking on "Yes". The aim here is to show, first, that the acceptability always follows the PA, and that the PA computation is enough to explain all type the dilemmas, and second, that the Reaction Times of both type of trials is similar.

The special dilemmas are as follows:

- **Scene 11**

1. You work in a major supermarket and you have just caught someone stealing. You decide to give her away to your superior.
2. She happens to be your neighbor. She is from a family of 6 and none of the adults have a permanent job.
3. This person has been stealing food for a week and has been hiding it under her coat.
4. By giving her away to your boss, there is a 90% chance that you will get an 8% salary increase.

The PA and Acc ratings per trial are shown in figure (48):

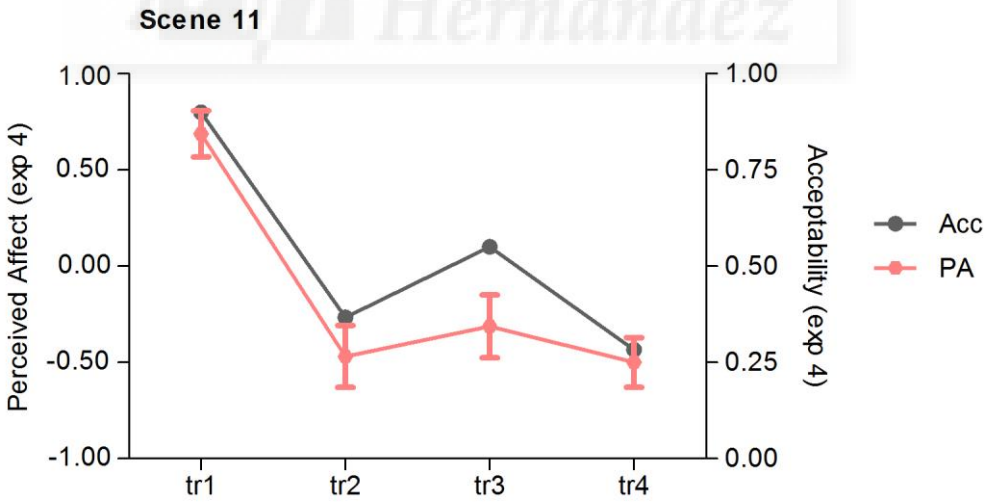


Figure 48. **Special Dilemma 1.** Relation between the average of PA (exp 4) and the Acc ratings (exp 4) n= 4 (trials).

- **Scene 12**

1. You love music. You mainly buy records by bands with limited financial means. So that their music reaches a larger audience, you decide to post it on-line.
2. Because of this behavior, these kinds of bands miss out on making large amounts of money every month.
3. You do not ask for these bands' permission in order to upload their songs to Internet.
4. In the long run, because of your actions, many people get to know these kinds of bands and therefore buy their records.

The PA and Acc ratings per trial are shown in figure (49):

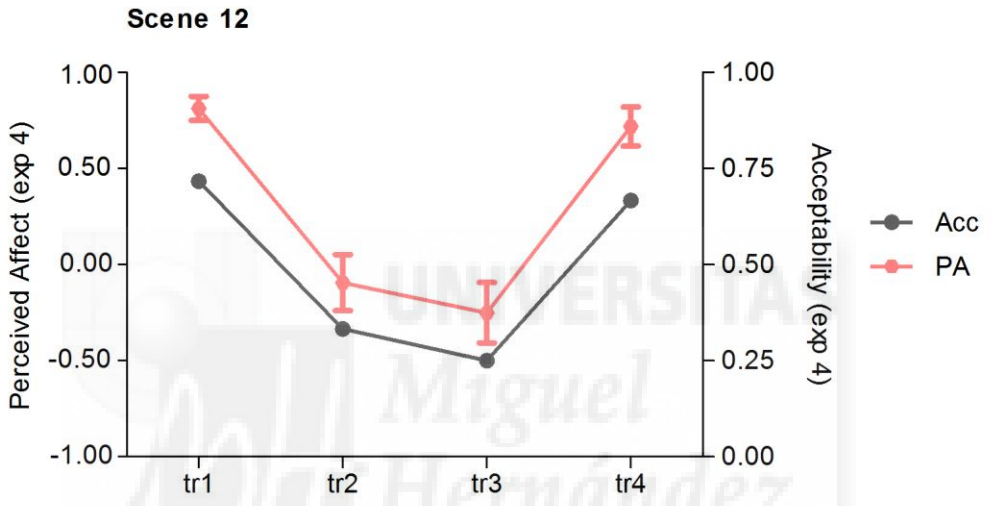


Figure 49. **Special Dilemma 2.** Relation between the average of PA (exp 4) and the Acc ratings (exp 4) n= 4 (trials).

- **Scene 13**

1. A girl decides to lend money to his best friend.
2. The borrowed money was intended for buying a variety of drugs, one of them being cocaine.
3. The friend needed the drugs in order to reduce the anxiety he was experiencing. He felt as if he was about to go crazy.
4. By lending him the money for the drugs, it was 80 % less probable that he would commit a knifepoint robbery.

The PA and Acc ratings per trial are shown in figure (50):

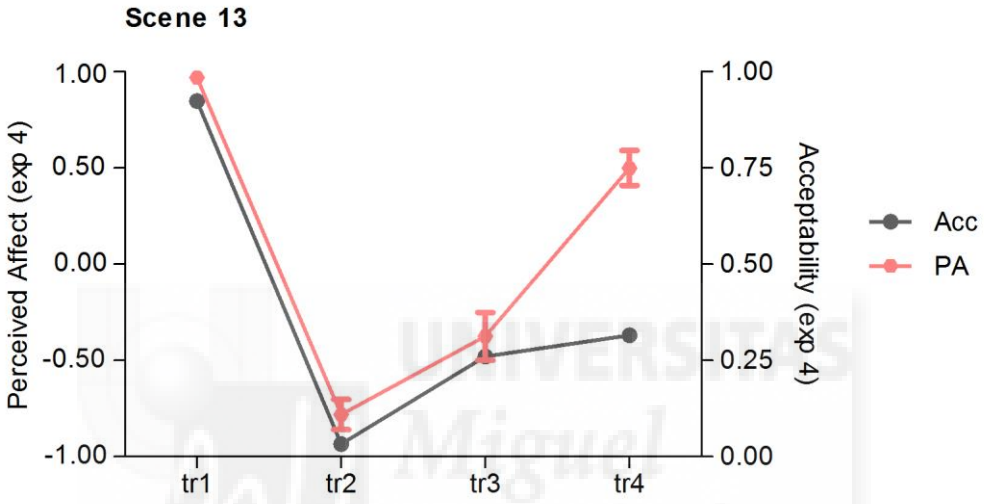


Figure 50. **Special Dilemma 3.** Relation between the average of PA (exp 4) and the Acc ratings (exp 4) n= 4 (trials).

- **Scene 14**

1. You are the surgeon responsible for carrying out an operation. Because you were distracted, you commit a serious mistake and the girl you were operating on suffers a serious injury.
2. The little girl had been brought to the hospital because of a blood flow problem to her brain. The surgery that was carried out was tremendously complicated.
3. You and your team mistook the syringe with glue for the syringe with the contrasting agent and administered glue, causing a fatal reaction to occur in her body.
4. Due to the health system, you are required to work 2 night shifts per month. When you were carrying out the operation you had already been working for 40 hours straight.

The PA and Acc ratings per trial are shown in figure (51):

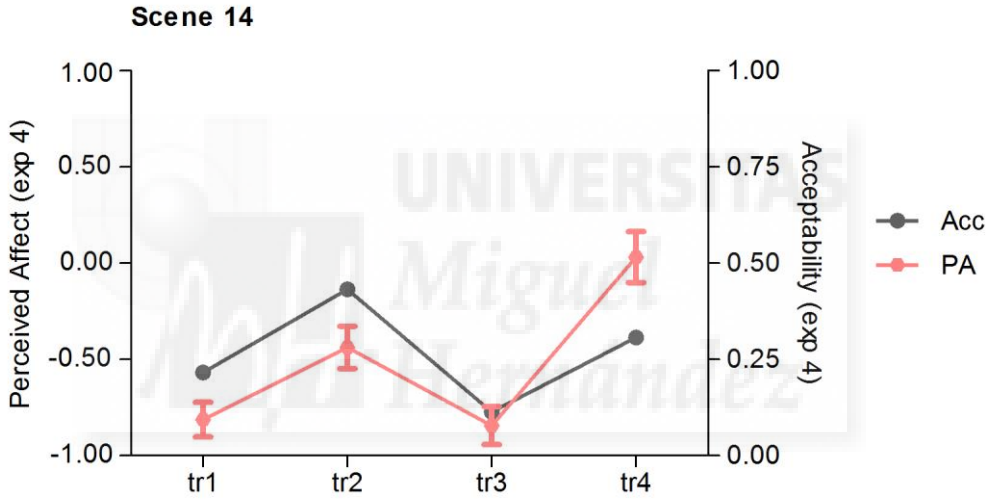


Figure 51. **Special Dilemma 4.** Relation between the average of PA (exp 4) and the Acc ratings (exp 4) n= 4 (trials).

- **Scene 15**

1. You are part of the medical staff of the Hospital of New Orleans and you have decided to leave your job before the arrival of Hurricane Katrina.
2. Your priority is to save your life.
3. Because you leave your job, the patients that you are in charge of will not survive, especially the five children with special needs that depend on you for their survival.
4. If you had opted for staying at the hospital to take care of them, they would have had a 25% chance of surviving.

The PA and Acc ratings per trial are shown in figure (52):

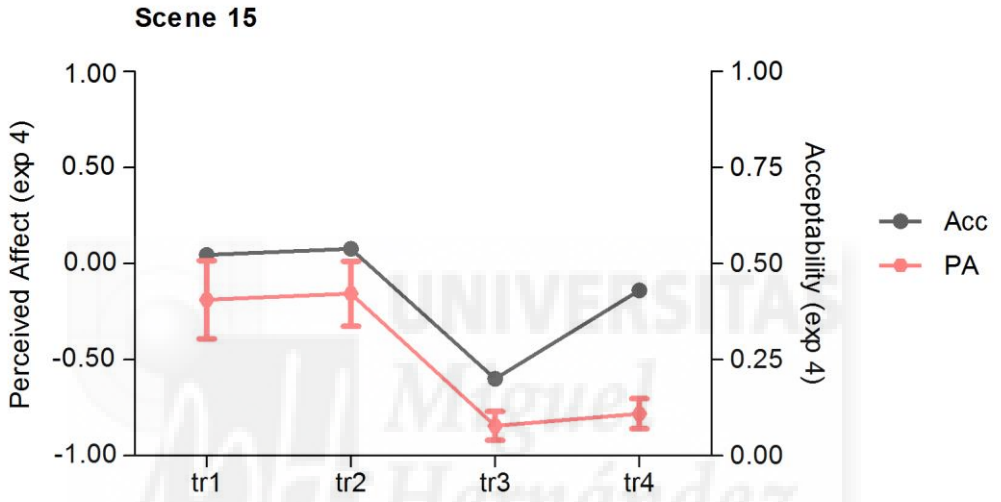


Figure 52. **Special Dilemma 5.** Relation between the average of PA (exp 4) and the Acc ratings (exp 4) n= 40 (trials).

We have correlated the average of the PA's trials with the proportions of the Acc of each trial, both obtained in experiment 4 (figure 53). Data shows PA had had a large impact on acceptability. The findings confirm there was a very strong correlation between both variables (r_s= 0.7570 n = 20 (trials), p< 0.0001).

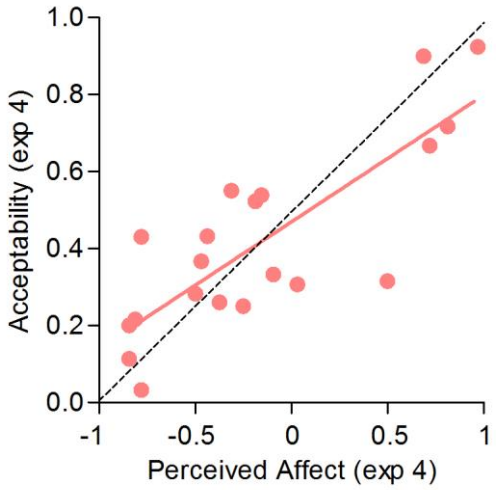


Figure 53. **Relation between PA & Moral Acceptability in the 5 Special Dilemmas**. Cross correlation graph between the average of PA (exp 4) and the Acc ratings (exp 4) of the 5 special dilemmas. The association was significant [r_s= 0.7570; n = 20 (trials); p< 0.0001].

Figure (54) shows the average RT of Deontological (M= 2.883, SEM= 0.181) and Contextual trials (M= 3.170, SEM= 0.151) respectively in the 5 special dilemmas. We have conducted Mann-Whitney test for testing whether the DR were faster than CR (U= 43678; p= 0.0972). We did not find statistics RT differences between DR and CR.

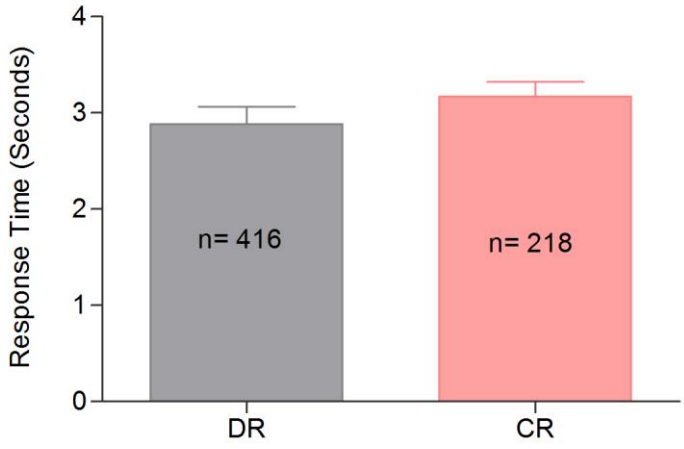


Figure 54. **RT differences between CR and DR in 5 Special Dilemmas.** Mean RTs (± SEM) by response type in experiment 3. A Mann-Whitney did not revealed statistical differences between Deontological (n= 218) or Contextual (n= 416) responses [U= 43678; p> 0.05].

Appendix XII- Drift Diffusion Model

Once the thesis was finished, we decided to test if using a single process model we could predict the RTs' and the proportion of both contextual and deontological responses.

Drift Diffusion Models (DDM) are a class of simple and powerful models originally developed to cover response times for simple binary decisions in perceptual discrimination tasks (Ratcliff & McKoon, 2008). The underlying assumption is that the brain extracts, per time unit, a constant piece of evidence from the stimulus (drift) which is disturbed by noise (diffusion) and subsequently accumulates these over time. This accumulation stops once enough evidence has been sampled and a decision is made. DDM have been used successfully to quantitatively analyse behavioral data, i.e., reaction times and accuracy (Ratcliff & McKoon, 2008; Crockett, Nurth-Nelson, Siegel, Dayan, & Dolan, 2014).

To perform the analysis trials were divided depending on the type of response, deontological (labelled as correct) or contextual (labelled as incorrect).

The standard DDM incorporates the following four parameters (figure 55; Wabersich, & Vandekerckhove, 2014):

1. The boundary separation α which defines the distance between the two boundaries, with the first boundary being at 0 and the second being at α

2. The bias parameter β defining the relative starting point of the diffusion process between the two boundaries (the absolute starting point z0 can be obtained by $z0=\beta \cdot \alpha$)

3. The drift rate parameter δ which captures the tendency of the diffusion process to drift towards the upper boundary

4. The non-decision time τ that defines the time that passes without any diffusion process going on (e.g., this incorporates the time needed to encode a stimulus and to execute a motor response).

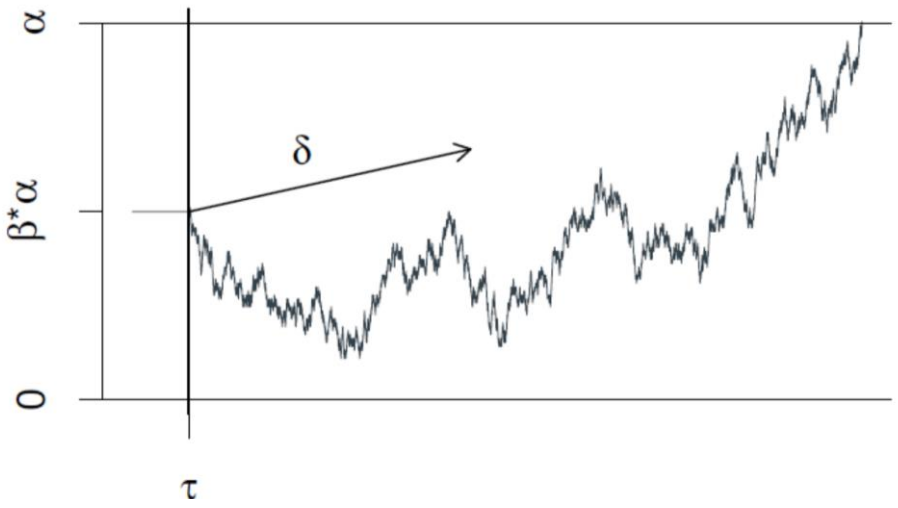


Figure 55. **Parameters of Drift Diffusion Model.**

According to the DDM (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006), the error rate is given by

$$ER = \frac{1}{1 + e^{\frac{2Az}{c^2}}}$$

And the mean decision time is given by

$$DT = \frac{z}{A} tanh\left(\frac{Az}{c^2}\right)$$

To get the mean RT we need to add the non-decision time τ

$$RT = DT + \tau$$

The results presented here show that the RTs' and proportion of both deontological and contextual responses can be modelled as a diffusion process (figure 56).

The estimated parameters are shown in table 43:

Table (43)
DDM Parameters

	M	SD
α	3.46	0.0367
β	0.512	0.00709
δ	0.307	0.0148
τ*	0.3	0.0206

* τ was fixated at 0.3 (Verdonck, & Tuerlinckx, 2016)

Empirical (Proportion Error, PE) and estimated (Error Rate, ER) contextual (incorrect) responses are shown in table 44:

Table (44)
DDM: Predicted and Estimated Contextual Responses

PE	M ER	SD ER	pvER
0.246	0.251	0.0104	0.618

PE is the empirical "proportion error", i.e. the fraction of contextual responses. M ER and SD ER, are the means and standard deviations of the fitted Error Rate. pvER is the p-value for testing the hypothesis Ho: PE == m ER

Empirical and estimated response time (RT vs RRT) are shown in table 45:

Table (45)
DDM: Predicted and Estimated RTs'

RT	M RRT	SD RRT	pvRT
3.03	3.16	0.0951	0.161

RT is the empirical response time, i.e. the mean value of all RTs. M RRT and SD RTT, are the means and standard deviations of the fitted Response Times. pvRT is the p-value for testing the hypothesis Ho: RT == m RTT

The parameters have been obtained with the R function "optim", and the SD with the Hessian. The SD of any function of the parameters have been obtained by simulating a gaussian distribution of the parameters and computing the resulting SD of the function distribution.

The P-value was calculated based on the Kolmogorov-Smirnov statistic test to evaluate if both the predicted and empirical data, follow the same distribution. The null hypothesis will be accepted for all values of α less than the P-value. We did not find statistical differences between both distributions (K-S test p< 0.01).

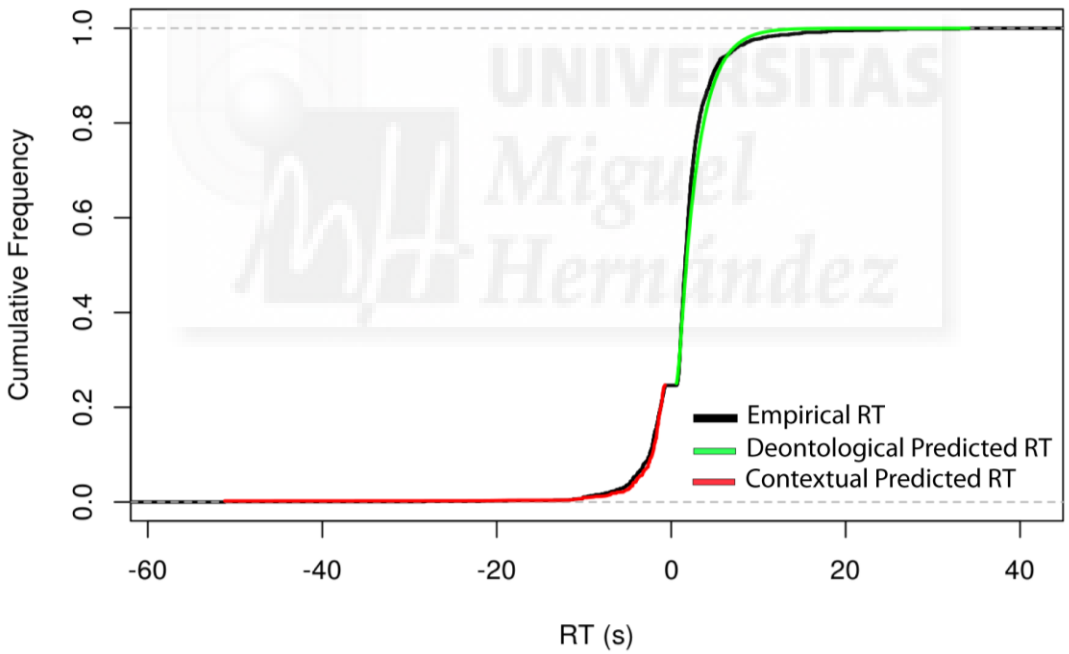
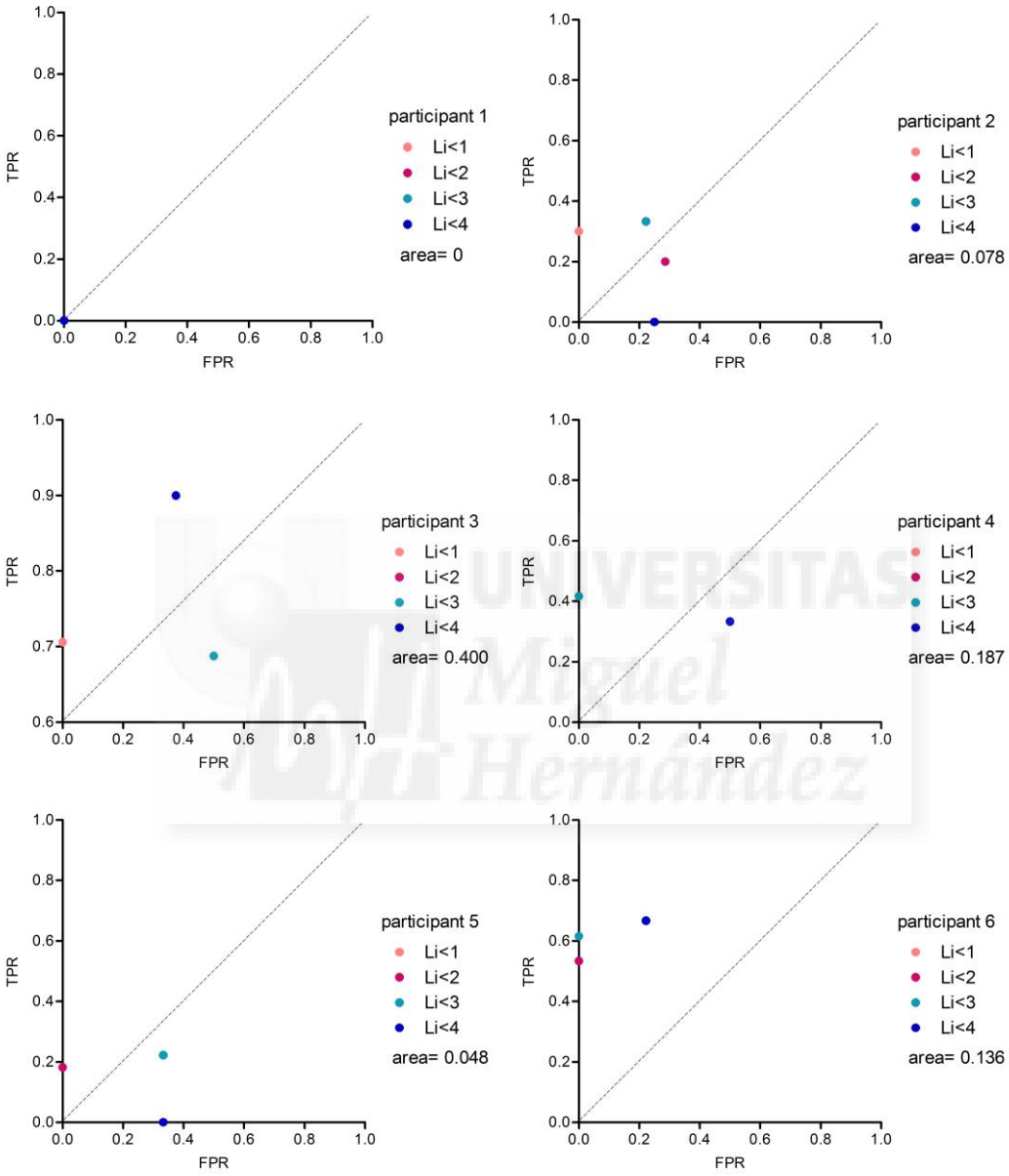
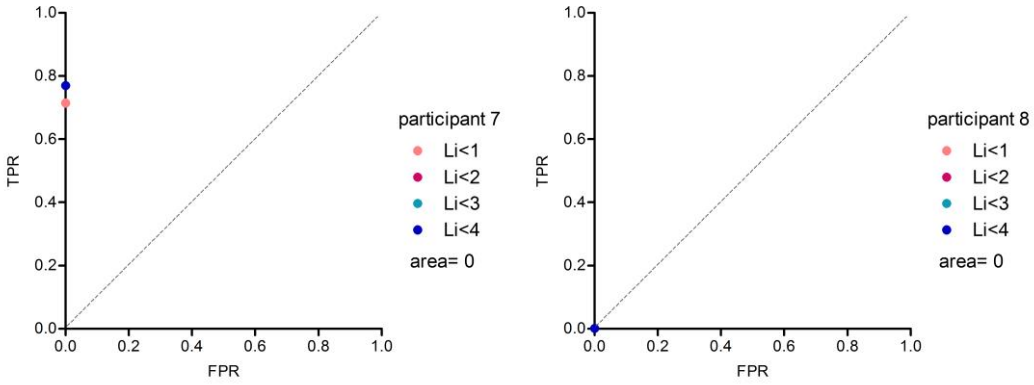


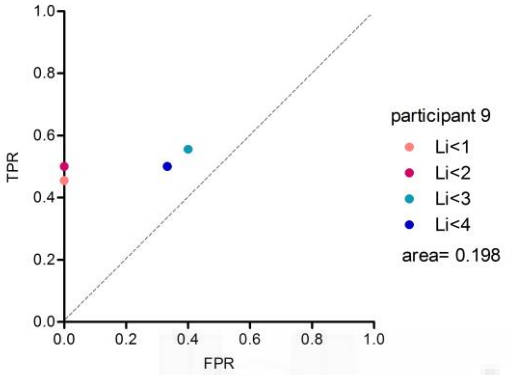
Figure 56. **Drift Diffusion Model.** The graph shows the cumulative frequency of a response at each RT. function. The black line represents the empirical data, where the Contextual responses were represented with negative time. Red line shows the predicted contextual RTs and green line the Deontological predicted RTs.

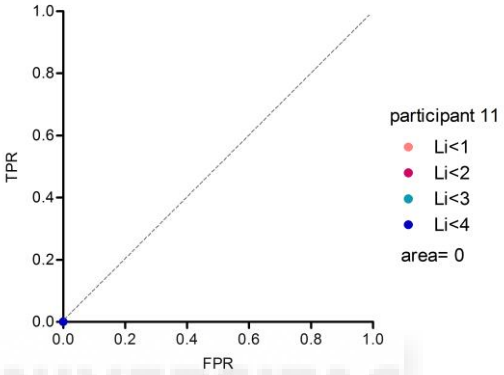
Apendix XIII- Individual ROC Space

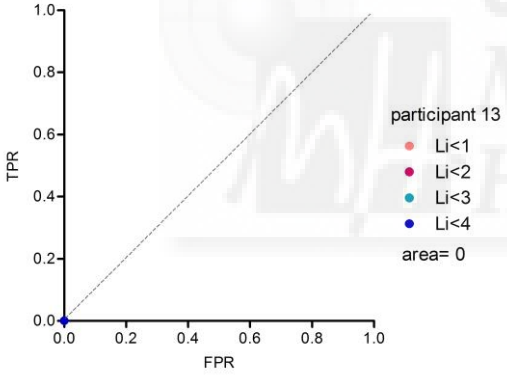
We calculated a ROC function for every single subject (figure 55). Each point plots the false alarm rate on the x-axes against the true positive rate on the y-axes for a given confidence criterion (Fleming & Lau, 2014).

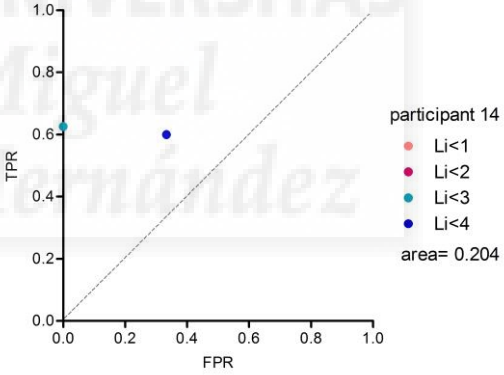


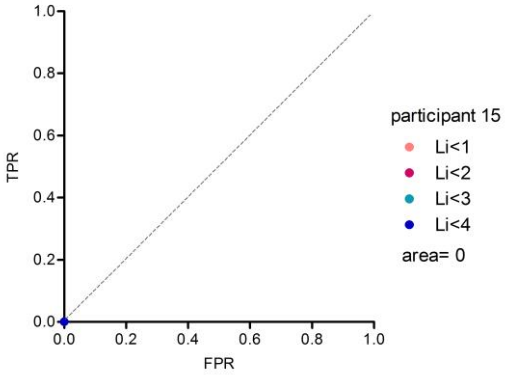


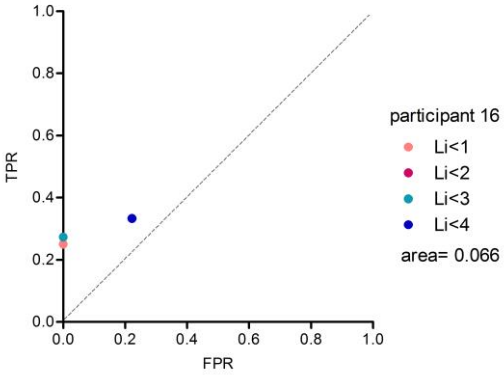


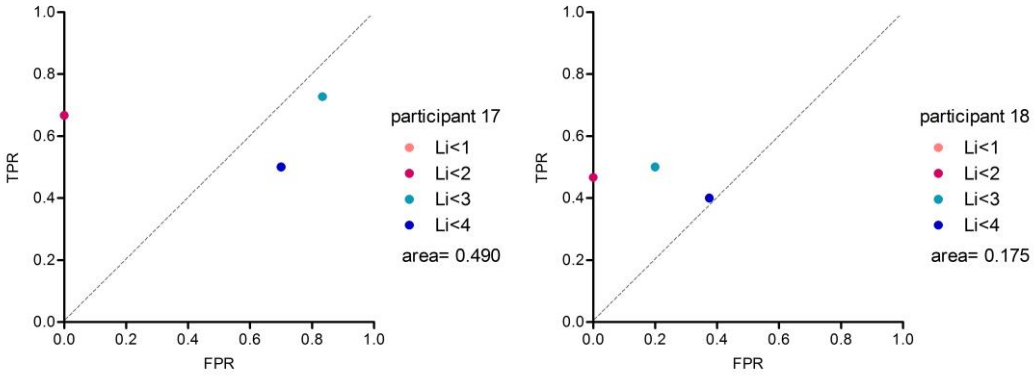


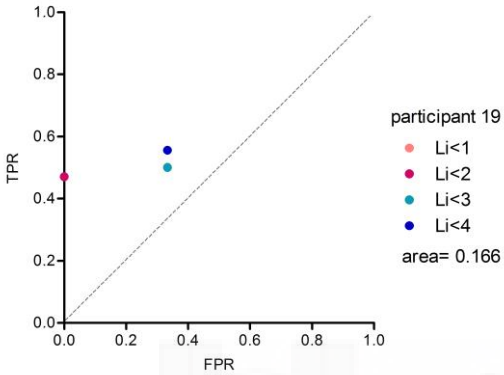


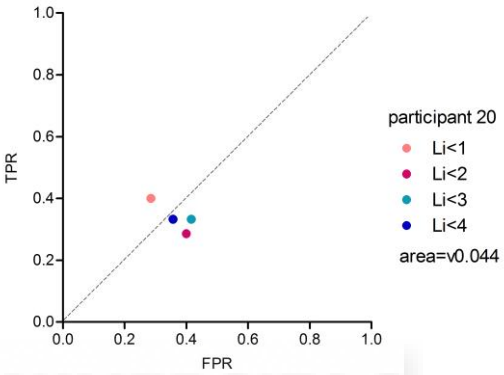


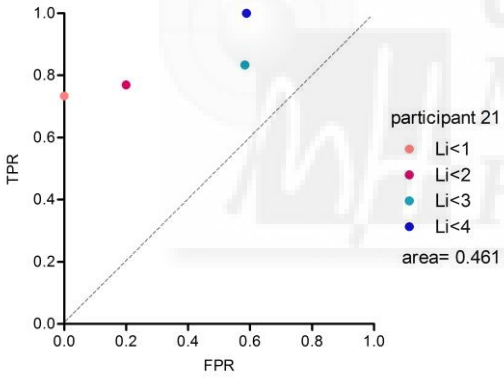


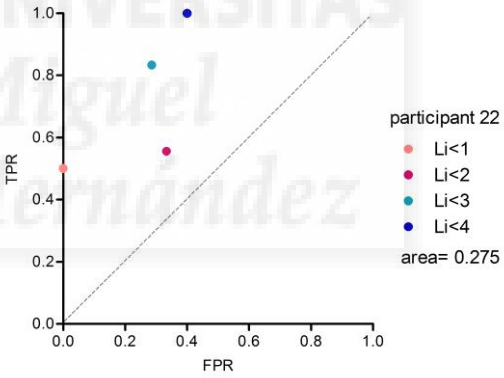


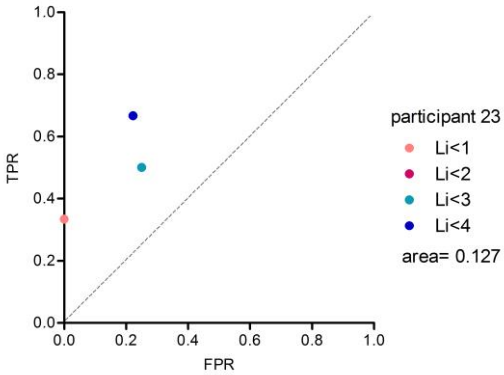


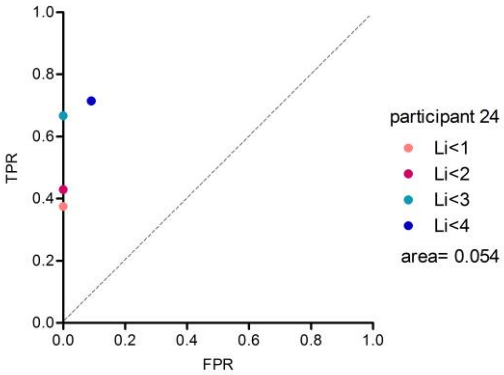


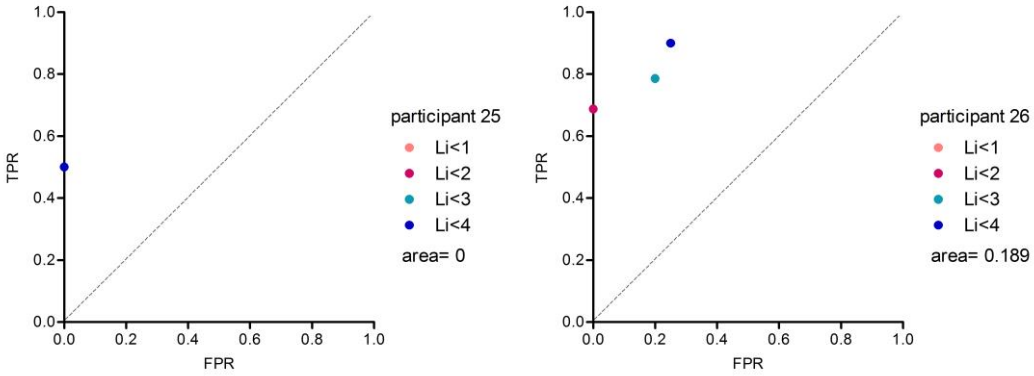


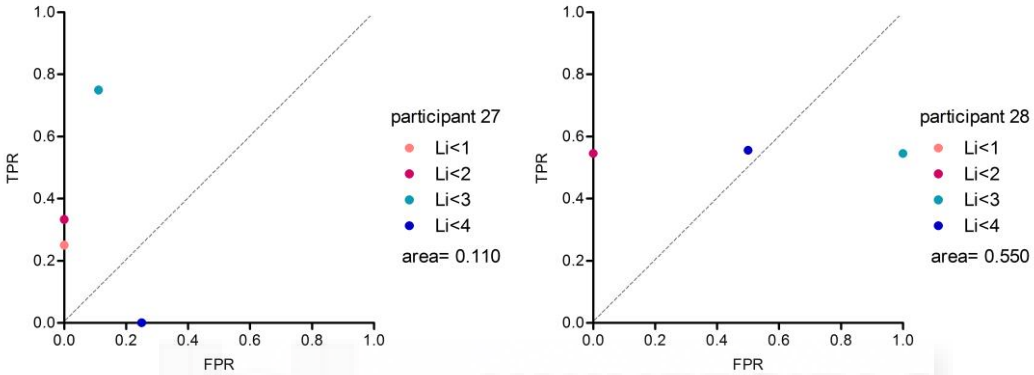


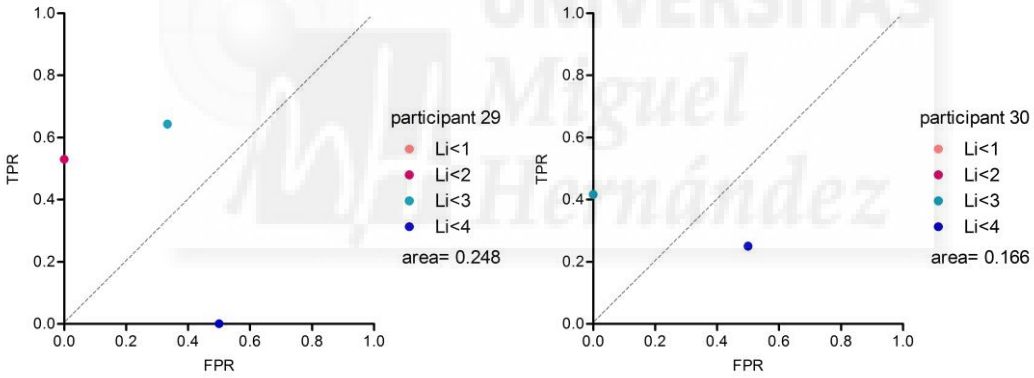


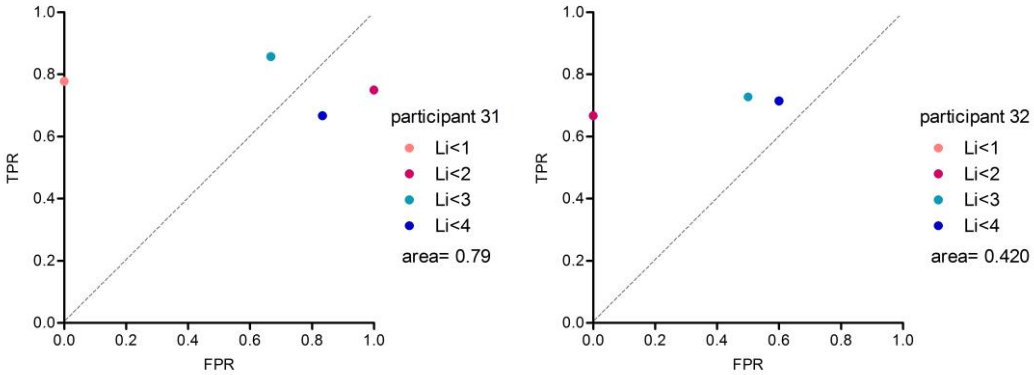












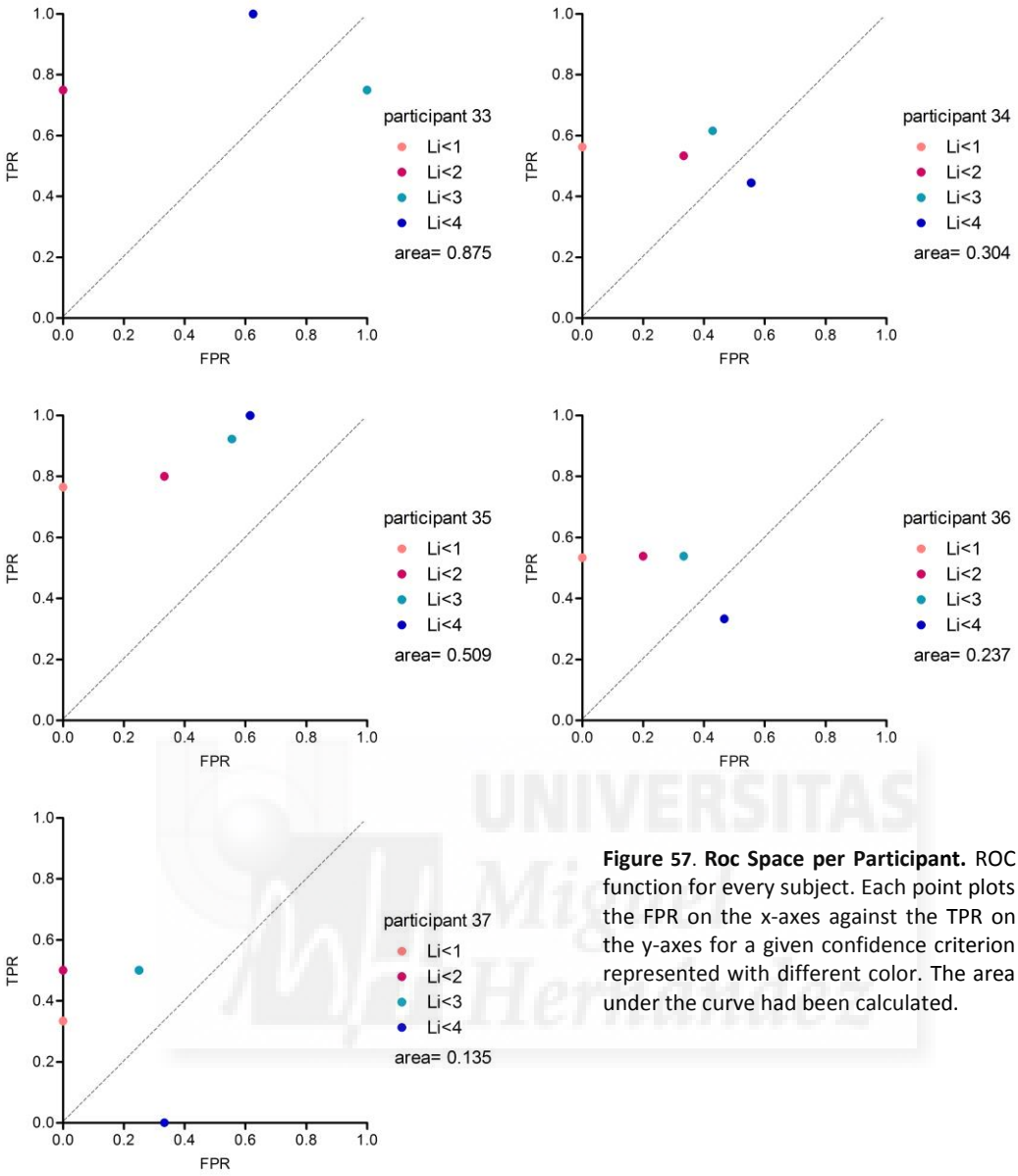
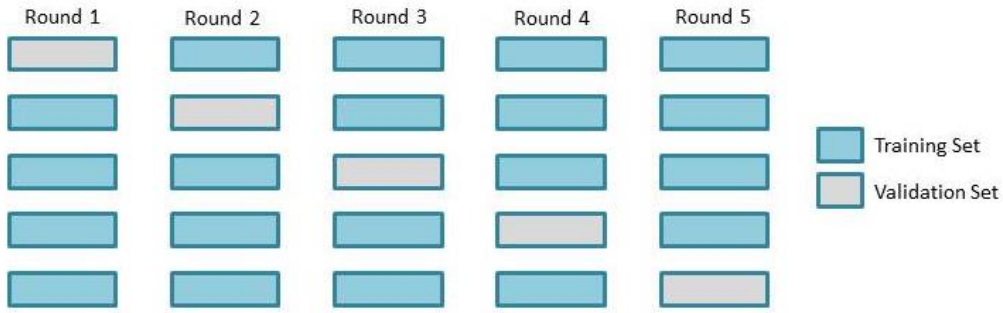


Figure 57. **Roc Space per Participant.** ROC function for every subject. Each point plots the FPR on the x-axes against the TPR on the y-axes for a given confidence criterion represented with different color. The area under the curve had been calculated.

Appendix XIV- Cross Validation

In order to evaluate the capacity of the Perceived Affect for obtaining the acceptability we have used the cross-validation tool to get the accuracy of the prediction. In this procedure, we have, firstly, sorted the data randomly and, secondly, it was divided into k (k=5) folds. Then data was run 'k' rounds of cross-validation. In each round, one of the folds is used for validation, and the remaining folds for training. After training the classifier, is needed to measure its accuracy on the validation data (figure 56). Finally, we have averaged the accuracy over the k rounds to get a final cross-validation accuracy. The accuracy numbers are shown in the figure 56.



Validation Accuracy for each round:

76.97% 76.28% 76.60% 77.48% 77.04%

General Validation Accuracy: 76.87%

Figure 58. **5 Fold Cross Validation.** The data set is divided into 5 portions or "folds". One fold is designated as the validation set, while the remaining 4 folds are all combined and used for training. The validation accuracy is computed for each of the ten validation sets, and averaged to get a final cross-validation accuracy. The accuracy numbers for each round and the general accuracy are shown below the picture.

Annexes

1. Consent Form
2. Experiment 1
3. Moral Dilemmas
4. Experiment 2
5. Control Questions
6. Recording Example Experiment 3
7. Experiment 4
8. Experimental Data

Annex 1: Consent Form

ID number:
Gender:
Age:
Level of studies:
Right handed/left handed:
Visual problems:
Religion and practice:

The aim of this research is to understand how the brain carries out. We need your participation to evaluate the morality or otherwise of each of the stories that we will present below.

The data obtained from your participation will be used only for scientific purposes and in no case make public your personal data, always ensuring full confidentiality of data and strict compliance of privilege in the use and management of information developed. If, even deciding to participate and consent collaboration initially, at some point in the intervention you want to stop participating, please notify us and, thus, you may leave the investigation without any compromise.

DECLARATION OF CONSENT

I, Mr / Mrs. ..... have read the informed consent document that has been given me, I understood the explanations provided therein about the purpose of research and why my participation is required, and I was able to resolve all doubts and questions I have raised the matter. I also understand that, at any time without giving any explanation, I can revoke the consent which I now offer. I have also been informed that my personal data will be protected and will be used only for scientific purposes. Taking all this into consideration and on such conditions CONSENT participate in research and that data arising out of my participation will be used to meet the objectives specified in the document.

In Leicester, a ..... of 2014

Signed:

UNIVERSITAS
Miguel
Hernández

Annex 2: Experiment 1

ID:
Sexo:
Edad:
Nivel de estudios:
Diestro/zurdo:
Religión/practicante:

El objetivo de la presente investigación es conocer el modo en que el cerebro lleva a cabo la toma de decisiones ante estímulos morales. Requerimos de su participación para que evalúe la moralidad o no de cada una de las historias que le presentaremos a continuación.

Los datos que se obtengan de su participación serán utilizados únicamente con fines científicos y en ningún caso se harán públicos sus datos personales, siempre garantizando la plena confidencialidad de los datos y el riguroso cumplimiento del secreto profesional en el uso y manejo de la información obtenida.

Yo, Don/Dña…….. he leído el documento de consentimiento informado que me ha sido entregado, he comprendido las explicaciones en él facilitadas acerca del objetivo de la investigación y el motivo por el cual se requiere mi participación, y he podido resolver todas las dudas y preguntas que he planteado al respecto.

También comprendo que, en cualquier momento y sin necesidad de dar ninguna explicación, puedo revocar el consentimiento que ahora presento. También he sido informado/a de que mis datos personales serán protegidos y serán utilizados únicamente con fines científicos.

Tomando todo ello en consideración y en tales condiciones, CONSIENTO participar en la investigación y que los datos que se deriven de mi participación sean utilizados para cubrir los objetivos especificados en el documento.

En San Juan de Alicante, a …... de …….. de 2015

Firmado:

A continuación le presentarán a una seria de personajes. Por favor, le pedimos que valore cuánto de bien le caen. Para ello, deberá indicarlo marcando una cruz en cada una de las casillas, siendo 1 muy poco y 5 mucho.

Muchas gracias por su colaboración.

	1	2	3	4	5
Futbolista	☐	☐	☐	☐	☐
Peluquero	☐	☐	☐	☐	☐
Trabajador ONG	☐	☐	☐	☐	☐
Jardinero	☐	☐	☐	☐	☐
Dentista	☐	☐	☐	☐	☐

	1	2	3	4	5
Soldado	☐	☐	☐	☐	☐
Alcalde	☐	☐	☐	☐	☐
Inmigrante argelino	☐	☐	☐	☐	☐
Cajera de cadena de supermercado	☐	☐	☐	☐	☐
Persona que descarga música	☐	☐	☐	☐	☐

	1	2	3	4	5
Amigo	☐	☐	☐	☐	☐
Enchufado	☐	☐	☐	☐	☐
Padre	☐	☐	☐	☐	☐
Hombre	☐	☐	☐	☐	☐
Cirujano	☐	☐	☐	☐	☐

	1	2	3	4	5
Madre	☐	☐	☐	☐	☐
Personal sanitario	☐	☐	☐	☐	☐
Mujer	☐	☐	☐	☐	☐
Profesor	☐	☐	☐	☐	☐

Annex 3: Moral Dilemmas

Dilemma 1

1. You are the commanding officer of a brigade in Afghanistan and you have decided to kill one of your men.
2. Your base has suddenly come under attack by nearby rebel militants. The soldier you just killed was seriously injured.
3. The soldier you just killed would not have been able to escape the rebels that were on their way.
4. If the rebels would have found him alive, there would have been a 75% chance that he would have endured serious torture and would have died a violent death.

Dilemma 2

1. You are the mayor of a town whose forest is going to be cut down. 1000 hectares of endemic vegetation are to be destroyed. You support this act.
2. Thousands of local citizens hold a demonstration outside the City Hall in order to protest against the measure. The forest is an emblematic symbol to the people of the region.
3. Behind the logging there is a whole real-estate and trade project.
4. 3,000 people will be employed for at least 3 years.

Dilemma 3

1. You are an Algerian citizen who decides to travel to Spain aboard a dinghy. You and your travel partner push eleven compatriots overboard without anybody else noticing in order to ensure you get to your destination safely.
2. You decide to push them overboard when you are a few hundred meters from the coast of Spain, just as the dinghy begins to sink. You do not know how to swim.
3. You are one of 48 people on the boat. If all 48 of you stay on the dinghy, it will end up sinking and only the good swimmers will survive.
4. You need to find a job in Spain in order to help your family in Algeria survive.

Dilemma 4

1. You just got a job in your line of academic training knowing that you are not the best candidate for the job since you got it through your contacts.
2. Given the current situation, this is the only way you can get a job.
3. Until now, you have been working on and off in the hospitality industry.
4. You have been waiting for an opportunity like this one since you graduated from University, that is, 5 years.

Dilemma 5

1. A man decides not to give away a member of the ETA terrorist group, knowing that he was planning an attack.
2. The terrorist begged him not to inform the Police.
3. The member of the terrorist group was his own son.
4. Not giving him away means that the likelihood of a future terrorist attack occurring increases by 60%.

Dilemma 6

1. An attractive, well-dressed, middle-aged man bursts into a pharmaceutical company and steals 3000€ worth of medication.
2. No one was willing to loan him the money in order for him to purchase the drugs.
3. His wife is seriously ill and needs to take a large amount of medication everyday.
4. By stealing the drugs, the probability of saving his wife's life increases by 50%

Dilemma 7

1. A mother lets her two-year-old baby die.
2. The baby was admitted to hospital in critical condition. He urgently needed to undergo surgery as well as a blood transfusion.
3. The mother did not authorize the blood transfusion since, being a Jehovah ’s Witness, her religion prohibits it.
4. There would have been an 80% chance of saving the baby had the transfusion been carried out.

Dilemma 8

1. The Supreme Court has reduced the nine-year sentence of a woman charged with burning a man alive, to four years.
2. The woman decided to set fire to a customer of the bar where she worked.
3. The customer she set fire to had raped her daughter a few months ago.
4. The customer she burnt alive was on temporary release from prison at the time of the incident and decided to go to the woman's workplace to make fun of her.

Dilemma 9

1. A teacher has hit his 15-year-old student
2. The 15-year-old student had come to school with the intention of hurting his fellow classmates.
3. This aggressive behavior towards his classmates has been going on for 5 months.

4. When the teacher hit him, the student was about to attack one of his classmates.

Dilemma 10

1. A mother lets her 11-month-old baby die.
2. The mother also has two more children, one aged 2 years old and one 4. They were all on their way home in the car when a strong storm broke causing a flood.
3. The mother managed to get all of them out of the car. Immediately, the flood carried the car away and she was able to find a tree to hold on to while holding onto her three children.
4. At a given moment, due to a lack of strength and the risk that the flood would take all four of them, she had to let go of the baby.

UNIVERSITAS
Miguel
Hernández

Annex 4: Experiment 2

A continuación escuchará una serie de historias. Cada una de las historias está compuesta por 4 escenas diferentes. Por favor, le pedimos que evalúe cuánto de sensible, agradable, amistoso, honesto y fiable le resulta el protagonista de cada una de ellas, es decir, cuánto de bien le cae. Para ello, deberá indicarlo marcando una cruz en cada una de las escenas, siendo **1 muy poco y 5 mucho**. Muchas gracias por su colaboración.

	1	2	3	4	5
1	1					
	2					
	3					
	4					
2	1					
	2					
	3					
	4					
3	1					
	2					
	3					
	4					
4	1					
	2					
	3					
	4					
5	1					
	2					
	3					
	4					
6	1					
	2					
	3					
	4					
7	1					
	2					
	3					
	4					
8	1					
	2					
	3					
	4					
9	1					
	2					
	3					
	4					
10	1					
	2					
	3					
	4					
11	1					
	2					
	3					
	4					
12	1					
	2					
	3					
	4					
13	1					
	2					
	3					
	4					
14	1					
	2					
	3					
	4					
15	1					
	2					
	3					
	4					

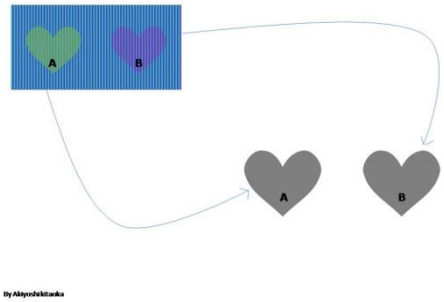
UNIVERSITAS
Miguel
Hernández

Annex 5: Control Questions

1. Visual Stimuli

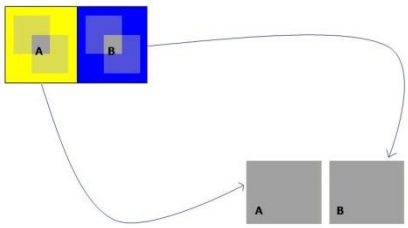
Illusion 1

- How do you perceive Hearts A and B?



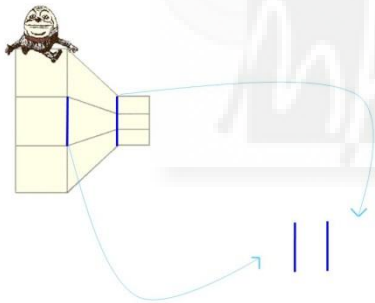
Illusion 2

- How do you perceive squares A and B?



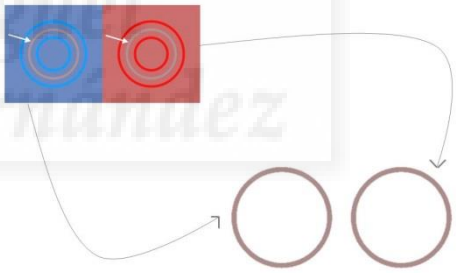
Illusion 3

- How do you perceive the blue lines?



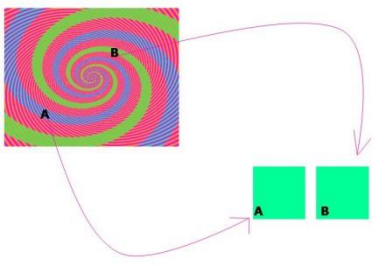
Illusion 4

- How do you perceive the circles?



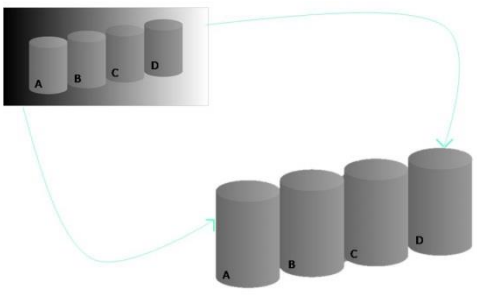
Illusion 5

- How do you perceive the bands A and B?



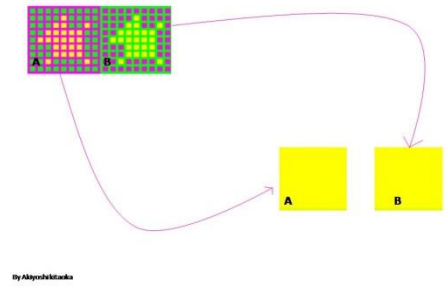
Illusion 6

- How do you perceive the cylinders A, B, C and D?



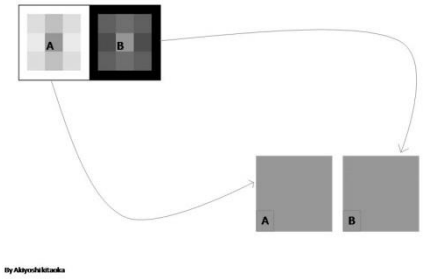
Illusion 7

- How do you perceive squares A and B?



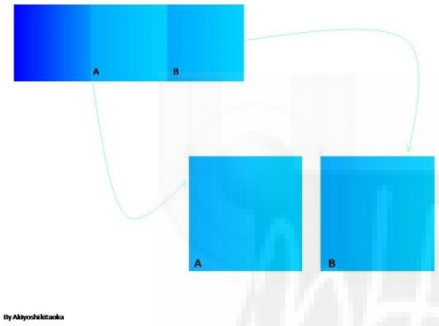
Illusion 8

- Do you perceive squares A and B as equals?



Illusion 9

- Do you perceive squares A and B as equals?



Note: all illusions except the number 3 are from Kitaoka, Illusion 3 is from Müller-Lyer, (1889).

2. Simple Questions

ORDER	CONTROL TYPE	QUESTION
1	Positive	Are you participating in an experiment?
2	Positive	Have you between 20 and 30 years?
3	Positive	Are you in a laboratory?
4	Negative	Are you in London?
5	Negative	Have you more than 80 years?
6	Negative	Are you standing?
7	False	Please answer a false answer, are you a guy?
8	False	Please answer a false answer, are you in Leicester?
9	False	Please answer a false answer, are you a student?

Annex 6: Moral Movie

Please see this movie for a real example:

https://www.youtube.com/watch?v=hbkT-7icgPg&feature=youtu.be

UNIVERSITAS
Miguel
Hernández

Annex 7: Experiment 4

Estimado estudiante,

En primer lugar muchas gracias por tu atención y colaboración.

Desde el Instituto de Neurociencias estamos lanzando un nuevo proyecto enmarcado en la Neurociencia Cognitiva, cuyo objetivo es proponer los mecanismos neuronales que subyacen a la toma de decisiones morales.

En la fase de la investigación en la que nos encontramos precisamos de tu colaboración. La tarea a realizar es muy sencilla, y no te llevará más de 10 minutos. Además, es totalmente anónima, no te pediremos ningún dato de identificación.

Al final de esta página encontrarás, en primer lugar, una plantilla que deberás rellenar. A continuación, en el proyector/pantalla de la sala van a aparecer una serie de dilemas morales. Cada dilema está formado por cuatro frases diferentes. Una vez que hayas escuchado cada una de los enunciados que componen el dilema, la tarea a realizar es la siguiente:

- Evaluar de cada una de las frases, cuánto de agradable, amistoso/a, honesto/a y fiable es el/la protagonista de la historia del 1 al 5, en resumen, cuánto de bien te cae, siendo 1 muy poco y 5 mucho.
- Evaluar de cada una de las historias cuánto de habilidoso/a, competente es el/la protagonista de la historia del 1 al 5, siendo 1 muy poco y 5 mucho.
- Evaluar de cada una de las historias presentadas si la acción te resulta moral o no. Para ello deberás señalar con una cruz la casilla SÍ o, por el contrario, la casilla NO.
- Evaluar de cada una de las historias con cuánto nivel de confianza estás respondiendo a la pregunta anterior (si la acción te parece moral o no) del 1 al 5, siendo 1 muy poco y 5 mucho.
- Indicar la emoción que sientes o en su defecto la que más se parezca.

Una vez hayas completado estas preguntas, se te presentará la siguiente parte del dilema o uno nuevo. Es muy importante que **NO MODIFIQUES** tus respuestas anteriores.

Si te encuentras con algún tipo de duda, por favor no dudes en preguntarnos. Los resultados de nuestro trabajo dependen de cómo lo rellenes. Muchísimas gracias.

Por favor marca con una cruz o completa los datos presentados a continuación. No olvides contestar a todas las cuestiones.

Edad ☐ Hombre ☐ Mujer ☐

Titulación ☐ Curso ☐ Universidad ☐

Sólo estudio ☐ Estudio y trabajo a tiempo parcial ☐ Estudio y trabajo a tiempo completo ☐

País de origen ☐ Ciudad de origen ☐ Ciudad actual ☐

Religión con la que ha sido educado ☐

Religión con la que se identifica en la actualidad ☐ ¿Es practicante? Sí ☐ No ☐

Colabora con alguna asociación, ONG, etc. Sí ☐ No ☐

Orientación política extrema izquierda ☐☐☐☐☐☐☐☐☐☐ extrema derecha

Número de hijos ☐

Según la información que has escuchado hasta ahora, deberás evaluar del 1 al 5 (siendo 1 muy poco y 5 mucho) cuánto de:

- **agradable, amistoso/a, honesto/a y fiable** te resulta el/la protagonista de la historia. En resumen, cuánto de bien te cae. Por favor, indícalo marcando una cruz en la primera fila de casillas.
- **habilidoso/a, competente** te resulta el/la protagonista de la historia. Por favor, indícalo marcando una cruz en la segunda fila de casillas.
- la acción que se te presenta **te resulta moral**. Por favor, indícalo marcando una cruz en la casilla SÍ (si la acción a tu juicio es moral) o por el contrario marca la casilla NO (en el caso de que no te resulte moral).
- con **cuánto nivel de confianza** estás respondiendo a la pregunta anterior. Por favor, indícalo marcando una cruz en la cuarta fila de casillas.
- qué **emoción** ha despertado en ti este dilema. En el caso de que no se corresponda con ninguna de las que aparecen a continuación, señala la que más se aproxime.

Número de dilema: 1

Personaje que evaluamos: responsable de la brigada

Parte 1

	Muy poco	**1**	**2**	**3**	**4**	**5**	**Muy**
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Parte 2

	Muy poco	**1**	**2**	**3**	**4**	**5**	**Muy**
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Parte 3

	Muy poco	**1**	**2**	**3**	**4**	**5**	**Muy**
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Parte 4

	Muy poco	**1**	**2**	**3**	**4**	**5**	**Muy**
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Según la información que has escuchado hasta ahora, deberás evaluar del 1 al 5 (siendo 1 muy poco y 5 mucho) cuánto de:

- **agradable, amistoso/a, honesto/a y fiable** te resulta el/la protagonista de la historia. En resumen, cuánto de bien te cae. Por favor, indícalo marcando una cruz en la primera fila de casillas.
- **habilidoso/a, competente** te resulta el/la protagonista de la historia. Por favor, indícalo marcando una cruz en la segunda fila de casillas.
- la acción que se te presenta **te resulta moral**. Por favor, indícalo marcando una cruz en la casilla SÍ (si la acción a tu juicio es moral) o por el contrario marca la casilla NO (en el caso de que no te resulte moral).
- con **cuánto nivel de confianza** estás respondiendo a la pregunta anterior. Por favor, indícalo marcando una cruz en la cuarta fila de casillas.
- qué **emoción** ha despertado en ti este dilema. En el caso de que no se corresponda con ninguna de las que aparecen a continuación, señala la que más se aproxime.

Número de dilema: 2

Personaje que evaluamos: alcalde

Parte 1

	Muy poco	1	2	3	4	5	Muy
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Parte 2

	Muy poco	1	2	3	4	5	Muy
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Parte 3

	Muy poco	1	2	3	4	5	Muy
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Parte 4

	Muy poco	1	2	3	4	5	Muy
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Según la información que has escuchado hasta ahora, deberás evaluar del 1 al 5 (siendo 1 muy poco y 5 mucho) cuánto de:

- **agradable, amistoso/a, honesto/a y fiable** te resulta el/la protagonista de la historia. En resumen, cuánto de bien te cae. Por favor, indícalo marcando una cruz en la primera fila de casillas.
- **habilidoso/a, competente** te resulta el/la protagonista de la historia. Por favor, indícalo marcando una cruz en la segunda fila de casillas.
- la acción que se te presenta **te resulta moral**. Por favor, indícalo marcando una cruz en la casilla SÍ (si la acción a tu juicio es moral) o por el contrario marca la casilla NO (en el caso de que no te resulte moral).
- con **cuánto nivel de confianza** estás respondiendo a la pregunta anterior. Por favor, indícalo marcando una cruz en la cuarta fila de casillas.
- qué **emoción** ha despertado en ti este dilema. En el caso de que no se corresponda con ninguna de las que aparecen a continuación, señala la que más se aproxime.

Número de dilema: 3

Personaje que evaluamos: argelino que empuja

Parte 1

	Muy poco	1	2	3	4	5	Muy
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Parte 2

	Muy poco	1	2	3	4	5	Muy
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Parte 3

	Muy poco	1	2	3	4	5	Muy
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Parte 4

	Muy poco	1	2	3	4	5	Muy
Agradable, amistoso, honesto y fiable		☐	☐	☐	☐	☐	
Habilidoso y competente		☐	☐	☐	☐	☐	
Moral		☐ SÍ		☐ NO			
Nivel de confianza de respuesta		☐	☐	☐	☐	☐	

Tristeza ☐ Miedo ☐ Alegría ☐ Ira ☐ Sorpresa ☐ Asco ☐

Annex 8: Experimental data

The following tables show the data obtained in the four experiments:

Table (39)

Results Experiment 1

Protagonist	M SA	SEM	n
1 Soldier	0.39	0.06	14
2 Mayor	0.36	0.07	14
3 Immigrant	0.48	0.05	14
4 “enchufado”	0.25	0.06	14
5 Man	0.68	0.04	14
6 Man	0.68	0.04	14
7 Mother	0.93	0.03	14
8 Woman	0.77	0.06	14
9 Teacher	0.73	0.04	14
10 Mother	0.93	0.03	14

Table (40)

Results Experiment 2

	tr	M PA	SEM	n		tr	M PA	SEM	n
Dilemma 1	1	-0.96	0.04	16	Dilemma 6	1	-0.96	0.04	16
	2	-0.5	0.1	16		2	-0.44	0.12	16
	3	0	0.14	16		3	0.6	0.1	16
	4	0.72	0.12	16		4	0.9	0.06	16
Dilemma 2	1	-1	0	16	Dilemma 7	1	-0.96	0.04	16
	2	-1	0	16		2	-0.9	0.06	16
	3	-0.88	0.08	16		3	-0.88	0.08	16
	4	0.1	0.12	16		4	-0.96	0.04	16
Dilemma 3	1	-0.74	0.1	16	Dilemma 8	1	-0.96	0.04	16
	2	-0.5	0.12	16		2	-0.96	0.04	16
	3	-0.28	0.12	16		3	0.34	0.12	16
	4	0.34	0.14	16		4	0.76	0.1	16
Dilemma 4	1	-0.16	0.14	16	Dilemma 9	1	-0.84	0.08	16
	2	-0.1	0.18	16		2	-0.34	0.16	16
	3	0.38	0.12	16		3	-0.12	0.18	16
	4	0.76	0.08	16		4	0.12	0.22	16
Dilemma 5	1	-0.96	0.04	16	Dilemma 10	1	-1	0	16
	2	-0.94	0.04	16		2	-0.68	0.08	16
	3	0.04	0.18	16		3	0.12	0.18	16
	4	-0.72	0.12	16		4	0.5	0.12	16

Table (41)

Results Experiment 3

	tr	Prop. Acc	n		tr	Acc	n
Dilemma 1	1	0	33	Dilemma 6	1	0	30
	2	0.19	33		2	0.03	30
	3	0.44	33		3	0.59	30
	4	0.75	33		4	0.69	30
Dilemma 2	1	0	30	Dilemma 7	1	0	35
	2	0	30		2	0.03	35
	3	0	30		3	0.15	35
	4	0.32	30		4	0.18	35
Dilemma 3	1	0.12	34	Dilemma 8	1	0.15	34
	2	0.24	34		2	0.06	34
	3	0.36	34		3	0.44	34
	4	0.48	34		4	0.50	34
Dilemma 4	1	0.43	30	Dilemma 9	1	0	32
	2	0.63	30		2	0.19	32
	3	0.63	30		3	0.23	32
	4	0.67	30		4	0.32	32
Dilemma 5	1	0	30	Dilemma 10	1	0	31
	2	0	30		2	0.03	31
	3	0.17	30		3	0.23	31
	4	0.03	30		4	0.43	31

Table (42)

Results Experiment 4

	tr	Prop. Acc	M PA	SEM	n		tr	Prop. Acc	M PA	SEM	n
Dilemma 1	1	0.07	-0.7	0.06	71	Dilemma 6	1	0.06	-0.76	0.04	99
	2	0.41	-0.22	0.08	71		2	0.30	-0.28	0.06	99
	3	0.41	-0.08	0.08	71		3	0.61	0.18	0.06	99
	4	0.65	0.24	0.08	71		4	0.77	0.4	0.06	99
Dilemma 2	1	0	-0.84	0.04	60	Dilemma 7	1	0	-0.96	0.02	60
	2	0.03	-0.8	0.04	60		2	0.05	-0.6	0.06	60
	3	0.07	-0.84	0.04	60		3	0.05	-0.88	0.04	60
	4	0.30	-0.28	0.08	60		4	0.07	-0.9	0.04	60
Dilemma 3	1	0.06	-0.84	0.04	71	Dilemma 8	1	0.14	-0.5	0.06	68
	2	0.08	-0.68	0.04	71		2	0.06	-0.72	0.06	68
	3	0.23	-0.5	0.06	71		3	0.45	0.08	0.08	68
	4	0.52	0	0.08	71		4	0.42	0.16	0.1	68
Dilemma 4	1	0.45	0.02	0.06	92	Dilemma 9	1	0.03	-0.88	0.04	66
	2	0.76	0.42	0.06	92		2	0.09	-0.64	0.06	66
	3	0.78	0.42	0.04	92		3	0.08	-0.56	0.06	66
	4	0.84	0.58	0.04	92		4	0.26	-0.18	0.08	66
Dilemma 5	1	0.01	-0.88	0.04	90	Dilemma 10	1	0.03	-0.82	0.04	66
	2	0	-0.82	0.04	90		2	0.27	-0.32	0.06	66
	3	0.12	-0.24	0.06	90		3	0.85	0.52	0.06	66
	4	0.02	-0.76	0.04	90		4	0.47	0.02	0.08	66

UNIVERSITAS
Miguel
Hernández