**UNIVERSIDAD DE LA LAGUNA**

**FACULTAD DE PSICOLOGIA**

**Departamento de Psicología Cognitiva, Social y Organizacional**

**Área de Psicología Básica**

**Cálculo intuitivo de probabilidad:**

**variables extensionales, intensionales y cognitivas**

**Tesis Doctoral** que presenta el licenciado en Psicología Gorka Navarrete García para la obtención del grado de Doctor.

Dirigida por **Carlos Santamaría Moreno**

**La Laguna 2008**

**AGRADECIMIENTOS**

Es interesante como la misma palabra puede tener significados tan dispares. Agradeces que la camarera traiga el café con una sonrisa, le agradeces a la mujer de tu vida su existencia, su presencia, que hace que el presente sea un lugar tan extraordinariamente hermoso, y le agradeces su aportación teórica a la persona a la que llevas cien páginas criticando.

No es difícil sentirse confuso, al fin y al cabo... vivimos en un mundo complejo, ya se sabe. En cualquier caso, me gustaría, con estas pocas líneas, plasmar, por burdamente que sea, mi más sentido agradecimiento a toda aquella buena gente sin la que uno sería probablemente otro. Muchos merecéis una mención especial, algunos varias. Pero habrá otras ocasiones mejores.

No obstante, hay una persona a la que no puedo obviar, y es que, sin ella, todo seguiría siendo un mero sueño, sombras en la pared de la cueva, interesantes, pero sombras al fin y al cabo. Las noches aun llenas de absurdos monstruos indestructibles, sudor frió y temblores. Pero no solo eso, Rut tiene al menos tanto merito como yo en lo que viene después por motivos que llevaría varias paginas explicar, en realidad varios años. Gracias por arrastrarme hasta este maravilloso lugar contigo.

Nada es como debiera, y uno le retorcería las entrañas al mismísimo diablo para que todo volviera a su lugar,... pero por desgracia la condición humana es la que es y lo que uno desea... solo eso, un solitario aleteo.

Pese a todo es bonito estar aquí, no hubiera sido posible sin vosotros.

Gracias.

Intuitive Probability Calculus: extensional, intensional and cognitive variables.

English Summary & Conclusions

**Index**

ENGLISH SUMMARY AND CONCLUSIONS 5

**INTRODUCTION 27**

ENLIGHTMENT AND HUMAN REASON 28

HEURISTICS AND BIASES 29

ECOLOGICAL RATIONALITY 31

ABOUT ANGRY DISPUTES 32

PROBABILITY CALCULUS, REPRESENTATION FORMAT AND

NATURAL FREQUENCIES 39

INFORMATION ADQUISITION AND REPRESENTATION 44

NUMBER REPRESENTATION 46

COMPUTATIONAL COMPLEXITY 46

ENTANGLEMENTS AND DISENTANGLEMENTS 48

SIMPLE AND BAYESIAN PROBABILITIES 51

GOALS AND HIPOTHESIS 51

**EXPERIMENTS 55**

**1. CONTEXTUAL FACTORS AND REPRESENTATION FORMATS 55**

1.1. SIMPLE-EVENTS PROBABILITIES I 57

Method 58

Participants and design 58

Procedure and materials 58

Results 60

1.2. SIMPLE-EVENTS PROBABILITIES II 63

Method 63

Participants and design 63

Procedure and materials 64

Results 65

1.3. SIMPLE PROBABILITIES AND CONTEXTUAL FACTORS 67

Method 68

Participants and design 68

Procedure and materials 69

Results 72

DISCUSSION: CONTEXTUAL FACTORS AND REPRESENTATION FORMAT

WITH SIMPLE PROBABILITIES 75

**2. BAYESIAN PROBABILITY CALCULUS, COMPLEXITY**

**AND REPRESENTATION FORMAT 81**

2.1. BASE-RATES MANIPULATION 87

Method 89

Participants and design 89

Procedure and materials 90

Results 92

2.2. REFERENCE CLASS AND REPRESENTATION FORMAT 95

Method 97

Participants and design 97

Procedure and materials 98

Results 104

2.3.2 Cognitive algorithms and CRT 108

2.3. ARITHMETIC COMPLEXITY AND REPRESENTATION FORMAT 110

Method 112

Participants and design 112

Procedure and materials 113

Results 117

2.4. ARITHMETIC COMPLEXITY AND ITS SUBTLETIES 120

Method 125

Participants and design 125

Procedure and materials 125

Results 128

DISCUSSION: COMPLEXITY AND REPRESENTATION FORMAT WITH BAYESIAN PROBABILITIES 129

**GENERAL DISCUSSION 135**

BIRD'S VIEW 135

ABOUT PALE LIZARDS 140

ARITHMETIC COMPLEXITY 142

CONTEXTUAL FACTORS 149

ADAPTATIONS AND FREQUENTIST HIPOTHESIS 151

IS THE MIND A GENERAL COMPUTATION SYSTEM OR A GROUP

OF ALGORITHMS OR SPECIALIZED MODULES? 158

CONCLUSION 159

**BIBLIOGRAPHY 163**

**Summary**

Everyday life is full of decisions that involve some kind of inference to the likelihood of events. We carry an umbrella if we think it is likely to rain, and seek medical advice if we fear we are ill. Since these decisions can sometimes be crucial, how good we are at estimating likelihood is a central psychological enquiry. For example, one might have to face the critical information presented in Problem 1 (Table 1).

The standard normative model for this kind of inference is Bayes’s Theorem. This estimates a posterior (or conditional) probability (e.g. of actually having breast cancer after a positive mammography) from the prior probability of each single event (the event of having breast cancer, and the event of having a positive mammography) and the reverse likelihood (having a positive mammography when actually having breast cancer). According to Bayes’s Theorem, in the described conditions, only 7.8% of the women that receive bad news do actually have breast cancer. However, more than 95% of physicians solving this problem estimated the probability to be over 70% (Eddy, 1982).

It is a well-established finding that this kind of Bayesian problem is better solved when presented in terms of absolute frequencies rather than single-event probabilities or percentages (e.g. Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). For example, when the breast cancer problem is phrased as in Problem 2, nearly half of the naive reasoners gave the Bayesian correct response (Gigerenzer & Hoffrage, 1995).

Table 1. Some Bayesian problems with different formats

Problem 1. “Normalized percentages” (from Eddy, 1982):

The probability of breast cancer is 1% for a woman at age forty who participates in routine screening. If a woman has breast cancer, the probability is 80% that she will have a positive mammography. If a woman does not have breast cancer, the probability is 9.6% that she will also have a positive mammography. A woman in this age group had a positive mammography in a routine screening. What is the probability that she actually has breast cancer? \_\_\_\_%

Problem 2. “Natural frequencies” (from Gigerenzer & Hoffrage, 1995):

10 out of every 1,000 women at age forty who participate in routine screening have breast cancer. 8 of every 10 women with breast cancer will get a positive mammography.

95 out of every 990 women without breast cancer will also get a positive mammography.

Here is a new representative sample of women at age forty who got a positive mammography in routine screening. How many of these women do you expect to actually have breast cancer? \_\_\_out of \_\_\_.

Problem 3. “Number of chances” (from Girotto and Gonzales (2001, Study 1):

A person who was tested had 4 chances out of 100 of having the infection.

3 of the 4 chances of having the infection were associated with a positive reaction to the test.

12 of the remaining 96 chances of not having the infection were also associated with a positive reaction to the test. Among 100 people who have a positive reaction to the test, the proportion that has the infection will be equal to \_\_\_ out of \_\_\_.

There has been considerable debate about the reasons for this apparent superiority of frequency formats. As Chase, Hertwig, and Gigerenzer (1998, p. 211) have put it: “*Human cognitive algorithms for this type of inference, if they exist, are most likely designed to operate on numerical information in the form in which humans have gathered it over evolutionary history: natural frequencies, that is, absolute frequencies that have not been normalized with respect to the base rates*”.

Contrary to the main prediction from the frequentist field, several articles have reported situations in which no ‘facilitation effect’ of frequencies over probabilities was found (Evans et al. 2000; Fiedler, Brinkmann, Betsch, & Wild, 2000; Girotto & Gonzales, 2001; Macchi, 2000; Macchi & Mosconi, 1998). For example, Girotto and Gonzales (2001, Study 1) compared a frequency-format problem to a probability problem expressed in ‘number of chances’ (Problem 3), which matched the structure of the frequency problem. They found no differences in the percentage of correct responses. Hoffrage and his colleagues (Hoffrage, Gigerenzer, Krauss, & Martignon, 2002) responded to this criticism by arguing that some of these authors used *normalized frequencies* instead of *natural frequencies* and that some of them used tricks to close the gap between probabilities and natural frequencies*.* For the frequentist approach, natural frequencies are the result of natural sampling, which is the way our species has encountered statistical information during millennia. Natural frequencies are not normalized, and this means the question should not refer to a new set of people with a positive reaction to the test, but to the same set described before. According to Gigerenzer and Hoffrage (1995), Bayesian reasoning with natural (i.e. absolute) frequencies involves fewer steps of mental computation than relative frequencies or probabilities (see also, Howson & Urbach, 1993; Macchi & Mosconi, 1998). Some studies have demonstrated that this sort of computational simplification might also be obtained by other means than natural frequencies (Evans et al. 2000; Fiedler et al. 2000; Macchi, 2000). However, it could be argued (as Hoffrage et al., 2002 have done) that in these experiments probability formats mimic natural sampling.

For the frequentist approach, natural frequencies are the cause of facilitation and not just one of the possible ways to simplify computation in Bayesian problems. If a module or algorithm of mind exists that has evolved to compute frequencies but not probabilities, the facilitation should occur in any kind of probability problem and its effect should be stronger than nonessential contextual factors. Otherwise, the frequency format could both facilitate and hinder reasoning. This point is not trivial. If natural frequencies are the cause of facilitation, it could be inferred that people are rational both in principle and in practice for this kind of reasoning, and errors should be attributed to an artifact of the presentation format. But if the effect of natural frequencies can be both positive and negative it cannot be claimed that people are “good intuitive statisticians” (Cosmides & Tooby, 1996), at least in practice.

According to evolutionary psychologists (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995), the human mind has evolved to process information in the format it is acquired in the real world (with examples ranging from “animal foraging to neural networks” (Gigerenzer & Hoffrage, 1995). The frequentist proposal establishes a misleading parallel between information gathering and information representation. However, it is not necessary to assume ‘a priori’ such a parallelism. An example can be found within the domain of neural networks, cited by Gigerenzer & Hoffrage 1995 as an illustration of sequential information gathering. Although neural networks do gather information in a sequential way, they represent that information in an amodal distributed probabilistic way (Bishop, 1997; Hinton, McClelland, & Rumelhart, 1986). In such a system, there would be an unnecessary computational cost associated with a frequentist representation. Furthermore, developmental and trans-cultural studies about number representation suggests less sophisticated systems as candidates for evolved core number systems (Dehaene, 1992, 1997; Feigenson, Dehaene, & Spelke, 2004). Although some controversy about the details remains, the proposed number systems are all decidedly too simple to deal with the relative complexity of a specialized system that deals with frequencies, as proposed by the frequentists. In any case, from the fact that the sequential nature of frequency formats can correspond to that of information acquisition through natural sampling (Gigerenzer & Hoffrage, 1995) it does not follow that the same correspondence can be found at the information representation and manipulation levels.

The whole representation format discussion is fairly complex and a number of confoundings and misinterpretations appeared throughout (Brase, 2002; Brase & Barbey, 2006; Girotto & Gonzales, 2002; Hoffrage et al., 2002; Johnson-Laird, Legrenzi, Girotto, Sonino-Legrenzi, & Caverni, 1999). There is little discussion about the role of factors such as semantic disambiguation (Hertwig & Gigerenzer, 1999; Nickerson, 1996) and clarification of the reference class (Gigerenzer, Hoffrage, & Kleinbölting, 1991). The critical issue remains the confounding between computational complexity and a special evolutionary status of frequencies. The trouble with using Bayesian problems to demonstrate the advantage of natural frequencies over other types of statistical data presentation is the severe problems found when attempting to disentangle the influence of mathematical complexity from data representation format.

Here we tried to throw some light on the discussion. We used simple and posterior probability experiments, and controlling the complexity by different means we showed how representation format is not an adequate discriminator. Things like reference class and arithmetic complexity can perfectly explain all the results found in the literature with a higher level of detail and prediction capability. Also, we showed, although less emphatically, how the evolutionary arguments defended by the frequentists are an unnecessary addition for the study of probability calculus.

The dissertation is divided in two parts. In the first part (experiments 1, 2 and 3) we used simple probability problems and in the second part (experiments 4, 5, 6 and 7) posterior probability problems. The experiments in the first part showed no relevant effect related to the representation format (natural frequencies/probabilities) and a strong and significant effect of a contextual factor as *Instructions-Response Matching*, explainable by computational complexity and Mental Models.

In the second part, again, computational complexity allowed predict the results, leaving an explanation based on representation format unnecessary and even contradictory with our findings. Furthermore, we found an important correlation between a cognitive capacity (Cognitive Reflection Test) and the accuracy in the classical Bayesian problems, which goes against the strong evolutionary claims found in the frequentist literature.

**Conclusions**

These results pinpoint some flaws in the main frequentist prediction. Not only did participants not perform better with natural frequencies than with probabilities but in fact, contextual factors such as I-R matching were found to be crucial for their performance and other factors as reference class and arithmetic complexity became clearly more necessary than representation format to explain the subjects responses. However, several ad hoc allegations could be made in defence of the frequentist approach:

A problem with the computational complexity hypothesis is the confounding we referred to in the introduction. If the only testable proposition of Gigerenzer and his colleagues is that when Bayesian problems are worded in a format that simplifies computation (with natural frequencies, ‘number of chances’, etc) the problems are better solved by the participants, there is not much to say to that. That multiplication is easier with Arabic numbers than with Roman numbers is not something new in the 21st century (see, Hoffrage et al., 2002). But while computational simplification seems the only testable proposition of their proposal, the frequentist papers are replete with other propositions concerning cognitive evolution. For example: “The evolutionary argument that cognitive algorithms were designed for frequency information, acquired through natural sampling, has implications for the computations an organism needs to perform when making Bayesian inferences.” (Gigerenzer & Hoffrage, 1995, p. 686). If our participants’ cognitive algorithms were “designed” for frequency information, it is hard to explain why they did not make the most of this when solving our frequency tasks. Moreover, we should keep in mind that they not only failed to take advantage of the frequency format; but their performance was significantly better with I-R matching (as compared to no I-R matching) in simple probability problems and we were able to create probability conditions where the subjects performed better with probabilities… it was a matter of the arithmetic complexity.

An allegation in defence of the frequentist approach could be that that simple frequencies (those reflecting simple probabilities) are not natural frequencies. Although this would then count against the evolutionary origin attributed to natural frequencies. Not many things about our cognitive system’s past can be taken for granted. However it is an analytic truth that each time one of our ancestors encountered the frequencies needed for a Bayesian probability, he or she should have encountered three simple frequencies.

When proposing something as an adaptation it is important to show that the alternative hypothesis cannot account for the current evidence (Andrews, Gangestad, & Matthews, 2003; Williams, 1966). It is also important to note the possibility of exaptation, particularly in the cognitive domain (Andrews et al., 2003) where, for example, neural networks present an option where a single learning mechanism can lie behind a number of different cognitive mechanisms (Kruschke, 1992).

The main concern with the frequentist hypothesis, as we have noted, is the difficulty to empirically disentangle it from its alternative, namely, the complexity hypothesis. Everyone agrees that frequency formats make computation easier in many situations (particularly in Bayesian problems). The empirical evidence shows that, in general, accuracy is higher in Bayesian problems stated in terms of natural frequencies. The problem lies in that these problems are also easier computationally, so the evidence can be reformulated to; in general, accuracy is higher in Bayesian problems when computational complexity is lower. Thus, the alternative hypothesis can account for the present evidence.

A number of factors can be (and have been) used to demonstrate how the complexity affects people’s performance (e.g. nested sets (Tversky & Kahneman, 1983; Sloman, Over, & Slovak, 2003)). Sadly the overlapping of predictions, some misunderstandings and subtleties about the nature of natural frequencies made the endeavour inconclusive. The only way to disentangle this is by finding a set of problems where, while controlling complexity, one varies representation format. Our results show that when complexity and representation format are untangled, complexity wins, shaped as *I-R matching* for simple probability problems and as number of calculus steps for Bayesian probability problems.

So long as natural frequencies are supposed to be the representation format favored by evolution there is a rather difficult explanation that the frequentist faction needs to give. How can a contextual factor, such as the matching between instructions and response have a greater influence than the representation format itself? Why subjects performance is better in some of our probability conditions?

A system of the kind presented by the frequentists should be based on an innate number representation system. Nevertheless so far, the only proposals of innate number systems are a lot less sophisticated than the frequentist hypothesis would require (Dehaene, 1992, 1997; Feigenson et al., 2004). There are, of course, some discrepancies about the subtleties of the different theoretical proposals. Yet there seems to be a widespread agreement concerning some central properties of number representation. Namely, that the number representations are imprecise, abstract and can be compared and combined by simple arithmetic operations (addition and subtraction; Spelke & Kinzler, 2007). Dehaene and his colleagues (Dehaene, 1992, 1997; Feigenson et al., 2004), for example, proposed two core systems of number. The first one *“yields a noisy representation of approximate number that captures the inter-relations between different numerosities…”* (Feigenson et al., 2004, p. 309). The second one is *“for precisely keeping track of small numbers of individual objects and for representing information about their continuous quantitative properties”* (Feigenson et al., 2004, p. 310). The critical claim is, in their own words:

*“The two core systems are limited in their representational power. Neither system supports concepts of fractions, square roots, negative numbers, or even exact integers. The construction of natural, rational and real numbers depends on arduous processes that are probably accessible only to educated humans in a subset of cultures, but which nevertheless are rooted in the two systems that are our current focus and that account for humans’ basic ‘number sense’”.* (Feigenson et al., 2004, p. 307)

It is hard for us, with such restrictions, to think of an innate module or system dedicated to I) keeping track of frequency of events (e.g. 196 of 623) and II) computing subsequent probabilities using frequencies. Finally, the fact that a basic number representation has been found in human infants, children, adults and non-human primates (Spelke & Kinzler, 2007) makes the possibility of an independent frequency module particularly tricky.

The *Ecological Rationality* view has proved valuable. They presented a “Simon’s scissors” approach to the study of human cognition (Todd and Gigerenzer 2001, 2007). The two equally vital blades representing the human mind and the contextual world; consequently indicating ‘the world’ as an important and unavoidable part of the equation. The problem arises when we take for granted the criticalness of an environmental observation. It is hard to avoid the temptation to use plausible evolutionary explanations when we look around and see frequencies are all over the place. But, as we said before, it is critical to show that no alternative can account for the evidence when proposing something as an adaptation (Andrews et al., 2003; Williams, 1966), which is not the case with the innate frequency module hypothesis. Regardless, the main frequentist prediction does not hold up to our experiments. We showed that, in simple probability problems, contextual factors (I-R matching) may have a significant influence where representation format does not.

Also, using classic Bayesian problems we showed how the single most important factor with the bigger explanation power is arithmetic complexity. We were able to create and destroy differences just by manipulating the number of calculus steps necessary to solve the problems, independently of representation format. Reference class also became an important factor, although again explainable by arithmetic complexity. In any case, the malleability of results when altering subtle cosmetic factors and the arithmetic complexity should be a warning against the strong frequentist (evolutionary) claim.

We could separate the contributions of this work in various dimensions. Fist, we saw how the difference between probabilities and natural frequencies can be explained through the differences in complexity of both formats. When we used simple probabilities, the differences disappeared. Any factor adding complexity to the resolution of a problem (with both simple and posterior probabilities) affects the performance of subjects independently of representation format. I-R matching gave us proof of that with simple probabilities.

Secondly, the influence of the presence of the total sample and the complementary set wasn’t significant in our results. That should force to soften certain positions or, at least remind about the necessity of using empirical findings to dismiss or criticize results.

Thirdly, we also found that general cognitive abilities seem to influence the proper resolution of the Bayesian problems, being the influence bigger on natural frequencies than relative probabilities, which certainly do not support the existence of a specialized module or algorithm for natural frequencies.

Finally, we saw that arithmetic or computational complexity, seen as calculation steps needed to solve the problem, it is an important tool for predicting the behavior of subjects and explain differences based on representation formats, classes of reference, structure the problem, and other factors.

**Cálculo intuitivo de probabilidad:**

**variables extensionales, intensionales y cognitivas**

Gorka Navarrete García

*"it's not a cry that you hear at night  
it's not somebody who's seen the light*

*it's a cold and it's a broken hallelujah"*

Para ti Dragoncita.

Índice

ENGLISH SUMMARY AND CONCLUSIONS 8

INTRODUCCION 27

Ilustración y razón humana 28

Heuristics and Biases 29

Ecological Rationality 31

Sobre airadas disputas 32

Cálculo de probabilidad, formato de representación y frecuencias naturales 39

Adquisición y representación de información 44

Representación del número 46

Complejidad computacional 46

Enredos y desenredos 48

Probabilidades simples y Bayesianas 51

Objetivos e Hipótesis 51

EXPERIMENTOS 55

1. Factores contextuales y formatos de representación. 55

1.1. Probabilidades de eventos simples I 57

Método 58

Participantes y diseño 58

Procedimiento y materiales 58

Resultados 60

1.2. Probabilidades de eventos simples II 63

Método 63

Participantes y diseño 63

Procedimiento y materiales 64

Resultados 65

1.3. Probabilidades simples y factores contextuales 67

Método 68

Participantes y diseño 68

Procedimiento y materiales 69

Resultados 72

Discusión: Factores contextuales y formatos de representación

con probabilidades simples. 75

2. Cálculo de probabilidad Bayesiano, complejidad y formatos de representación 81

2.1. Manipulación de probabilidades previas 87

Método 89

Participantes y diseño 89

Procedimiento y materiales 90

Resultados 92

2.2. Clases de referencia y formatos de representación 95

Método 97

Participantes y diseño 97

Procedimiento y materiales 98

Resultados 104

2.3.2 Algoritmos cognitivos y CRT 108

2.3. Complejidad aritmética y formatos de representación 110

Método 112

Participantes y diseño 112

Procedimiento y materiales 113

Resultados 117

2.4. Complejidad aritmética y sus matices 120

Método 125

Participantes y diseño 125

Procedimiento y materiales 125

Resultados 128

Discusión: Complejidad y formatos de representación

con probabilidades Bayesianas 129

DISCUSIÓN GENERAL 135

A vista de pájaro 135

Sobre pálidos lagartos 140

Complejidad Aritmética 142

Factores Contextuales 149

Adaptaciones e hipótesis frecuentista 151

¿Es la mente un sistema general de cómputo o un conjunto

de algoritmos o módulos especializados? 158

A modo de conclusión 159

Bibliografía 163

***“****I give three arguments in favor of this conclusion, and argue against all three.****”***

**Unknown**

# INTRODUCCION

Una de las discusiones más acaloradas de los últimos años dentro de la psicología cognitiva ha sido la relativa a la influencia que determinados formatos de representación podrían tener en la exploración estadística del mundo que nos rodea en general, y el cálculo de probabilidad en particular. Las ramificaciones de esta disputa han alcanzado otras disciplinas como la psicología evolutiva, la cognición matemática (*mathematical cognition*), el razonamiento y la teoría de la evolución, y la batalla se ha librado en materias como la percepción, memoria, representación, evolución, modularidad de la mente y estadística, entre otras.

Es éste uno de esos casos en los que no es necesario tratar de manera grandilocuente la profundidad de una disputa altamente especializada y carente de interés para los profanos, al fin y al cabo están en juego cosas como la estructura de la mente y la falibilidad humana. La polémica, que en sus capas superficiales versa sobre qué formatos de representación son mejores o peores para el cálculo de probabilidad, se entronca en varias de las discusiones más importantes de la historia de la psicología y de la filosofía. Desde la estructura de la mente humana, ¿es ésta una colección de módulos especializados o una herramienta de cómputo de carácter general?, hasta la misma naturaleza de la razón, ¿es el ser humano una criatura sumida en la confusión, vulnerable a numerosos sesgos de pensamiento y dada al error o, por el contrario, extraordinariamente bien adaptada al mundo exterior?, pasando por cuestiones tan trascendentes como el papel del entorno en la evolución tanto filogenética como ontogenética de la cognición.

### Ilustración y razón humana

Hubo un tiempo en el que de la visión dominante sobre la razón humana brotaba un entusiasmo casi ilimitado. Éramos por aquel entonces una especie dotada de magníficas capacidades sin parangón en el mundo animal, ni punto de referencia externo que pudiera hacernos dudar sobre lo extraordinario de nuestra condición. Según la visión ilustrada, la razón era de una precisión encomiable, perfecta. Tanto era esto así, que cuando aparecían discrepancias entre lo que dictaba la teoría de la probabilidad y el juicio de los hombres razonables, se cambiaban las teorías (Daston, 1980; Gigerenzer, 1991; Gigerenzer y Hoffrage, 1995; Hacking, 1975). Un ejemplo sería la *St. Petersburg paradox*, según la cual, dos jugadores (A y B) se enfrascan en un juego lanzando una moneda al aire. Si sale cara en el primer lanzamiento, B le da a A un ducado. Si la primera cara sale en el segundo lanzamiento, B paga dos ducados. Si la cara no sale hasta el lanzamiento n, B paga 2n-1. Según la definición estándar de esperanza matemática (*expectation*), ésta es infinita y, por lo tanto, A tendría que estar dispuesto a pagar una cantidad infinita para entrar en el juego. La paradoja radica en que nadie en su sano juicio pagaría una cantidad alta para entrar en un juego así, por mucho que diga la teoría matemática. Nicholas Bernoulli fue uno de los hombres sabios para los que lo que decía la teoría no tenía sentido, y su sobrino, Daniel Bernoulli publicó en 1738 una resolución que distinguía entre dos tipos de esperanza, una matemática y otra moral (Bernoulli, 1954). Lo destacado es que esta resolución basada en el sentido común de un “hombre sabio” fue considerada como una importante contribución a la teoría matemática de la probabilidad (Daston, 1980).

### Heuristics and Biases

Por desgracia, la realidad empezó a interponerse dando muestras ineludibles de falibilidad en los otrora bastiones inexpugnables del pensamiento de los hombres. Surgió de la explotación de la vulnerabilidad de la razón una línea de investigación extraordinariamente fructífera que llegó a dar a uno de sus máximos exponentes un premio Nobel, el primero para un psicólogo (al menos si no tenemos en cuenta a Herbert Simon). Esta tradición se conoce como *Heuristics and Biases* (heurísticos y sesgos) (Kahneman, Slovic, y Tversky, 1982; Tversky y Kahneman, 1974, 1983). Kahneman, el premio Nobel, y Tversky, que también hubiera recibido este premio de no ser por la curiosa costumbre de no concedérselo a los muertos, fueron sus paladines más destacados. Su visión se alejaba, por lo tanto, de la perfecta e impermeable razón ilustrada, prestando más atención a los límites de ésta. Se empezó a ver la mente como una máquina de procesamiento de información de propósito general, aunque con limitaciones (de memoria de trabajo, de serialidad, de potencia de cómputo, etc.). Surgieron evidencias de cosas como el sesgo de Exceso de Confianza (*overconfidence bias*), la Falacia de la Conjunción (*conjunction fallacy*), y la Desestimación de Probabilidades Previas (*base-rate neglect)*. Y como consecuencia, los modelos normativos sobre el comportamiento humano tuvieron que incluir modificaciones que los alejaban de la perfección matemática o estadística de antaño (véase por ejemplo el caso de la Prospect Theory; Kahneman y Tversky, 1979).

Aunque el objetivo principal era entender los procesos cognitivos que producían juicios tanto válidos como inválidos, el énfasis principal se puso sobre los sesgos, por aquello de llevar la contraria al paradigma dominante, porque era un campo muy bien abonado y por su utilidad para entender los procesos subyacentes a cosas como el razonamiento (Kahneman y Tversky, 1996). Pero los sesgos no bañaban de manera indiscriminada la cognición humana. Un interesante fenómeno, que aparentemente se producía al superar los límites de la razón, se daba en la forma de los llamados heurísticos. Éstos fueron descritos como atajos cognitivos usados para compensar las carencias del sistema cognitivo, y era principalmente al usarlos cuando surgían, a veces, los sesgos. Dicho de otra manera, la razón era precisa, siempre y cuando no se superaran sus límites. No obstante, si lo más accesible o representativo no fuera generalmente lo más probable, no existirían los heurísticos. Normalmente, los ejemplos que a uno le vienen antes a la cabeza pertenecen a las categorías más frecuentes, aunque esto no siempre es así, porque la memoria no se nutre únicamente de la frecuencia, también lo hace de la saliencia de los ejemplos, entre otras cosas, y eso da lugar a sesgos. A partir de ahí, la visión del “Homo Economicus”, ejemplo de racionalidad, se difuminaba hasta quedar apenas una caricatura, sesgada.

Precisamente a causa de este enfoque tan centrado en los errores y sesgos, hubo algunas voces críticas con la metodología de Kahneman y Tversky. No todo el mundo estaba de acuerdo con la visión de la mente humana que, de manera no del todo voluntaria, surgía de este enfoque (Kahneman y Tversky, 1996), como una herramienta salpicada de fallos (Cohen, 1981; Einhorn y Hogarth, 1981; Lopes, 1991).

### Ecological Rationality

Y claro, las cosas no quedaron así. Surgió como reacción a la visión de los heurísticos y sesgos (Gigerenzer, 1991, 1993; Gigerenzer, 1994; Gigerenzer, Hell, y Blank, 1988; Gigerenzer y Murray, 1987) un grupo que ponía el énfasis en un lugar diferente, la segunda hoja de la tijera de Simon, el mundo exterior. Para Herbert Simon, al estudiar la cognición humana carece de sentido tener en cuenta únicamente la mente. Igual que unas tijeras son apenas un trozo de metal inservible a falta de una de las hojas, ignorar el mundo que nos rodea hace fútil cualquier esfuerzo por entender los íntimos entresijos de la cognición, reflejo ésta, por lo tanto, de una sutil danza entre mente y entorno (Simon, 1990, pag. 7).

El cambio de énfasis que trajo consigo la *Ecological Rationality* (o Racionalidad Ecológica) significa algo más que un renovado interés en el mundo que nos rodea. Nace esta postura de una oposición poliédrica, volviendo a una forma matizada de aquella hermosa visión ilustrada de la que hablábamos. Los errores dejan de ser fruto de las limitaciones de la mente humana. El denso patrón de incorrecciones encontrado por los archienemigos partidarios de los heurísticos y sesgos surge de un fallo por parte de los investigadores al situar a los sujetos experimentales en situaciones alejadas de aquellas para las que la mente fue diseñada (Cohen, 1981), o simplemente por una excesiva rigidez al buscar que las respuestas se ajusten a modelos normativos erróneos o inadecuados para el problema usado (Gigerenzer, 1991; Gigerenzer y Hoffrage, 1995). Sería como introducir números para una suma en una calculadora (decimal) en clave binaria y además esperar que dé la respuesta en números romanos.

También cambió con ellos la visión de los heurísticos, que pasaron a ser una manera extraordinariamente eficiente de tratar con el mundo, desproveyéndolo de la complejidad habitual y convirtiéndolo en un juego de niños (Brighton, 2006; Gigerenzer y Brighton, 2007; Gigerenzer, Hoffrage, y Kleinbölting, 1991).

La *Ecological Rationality* disgregó la visión que tenía a la mente como una máquina de procesamiento de información de propósito general, para convertirla en un conjunto de piezas diseñadas por selección natural para tratar con distintos aspectos de la realidad. Así que las limitaciones que la mente pudiera tener pasaron a ser oportunidades. La historia filogenética nos ha moldeado a imagen y semejanza del entorno, supliendo las carencias con funciones adaptativas (a través de algoritmos o módulos especializados) que ofrecen en su lugar respuestas extraordinariamente ajustadas, casi mágicamente (Brase y Barbey, 2006).

Por lo tanto, el estudio del mundo como complemento al de la mente no es tan sólo un capricho, sino algo de gran utilidad, ya que las regularidades en su estructura nos permiten inferir aquellos dominios para los que existen adaptaciones cognitivas. Y no hablamos únicamente del mundo que nos rodea actualmente. El énfasis de la *Ecological Rationality* en aspectos evolutivos señala la importancia de tener en cuenta el mundo que nos ha rodeado a lo largo de nuestra historia evolutiva (Todd y Gigerenzer, 2000, 2002, 2007).

### Sobre airadas disputas

En el fondo, las visiones de los Heurísticos y Sesgosy la *Ecological Rationality* son más similares de lo que las disputas parecen mostrar. En realidad, muchas de las diferencias serían más de matiz. Para los primeros, los heurísticos existen porque en la mayoría de las ocasiones dan la respuesta correcta, solo que también dan lugar a sesgos. Su investigación se centró de manera prioritaria en los errores, como decíamos antes, por su utilidad para estudiar cómo funciona la mente, pero también, seguramente, porque la prevalencia de éstos era enorme. El matiz está en que, para los segundos, los heurísticos son una adaptación al medio, y fallan principalmente bajo dos condiciones. Cuando se saca al sujeto o a la habilidad cognitiva de su medio natural (mostrando información en forma de probabilidades en lugar de frecuencias naturales) y cuando se usan criterios normativos erróneos (no avisando explícitamente de que el muestreo es aleatorio y presuponerlo normativamente). Esto último fue una fuente muy importante de disputas.

Uno de los sesgos más famosos es la *desestimación de probabilidades previas*. Hay un amplio acuerdo sobre la extensión y robustez de su existencia (por ejemplo, Bar-Hillel, 1977; Bar-Hillel y Fischhoff, 1981; Landman y Manis, 1983), aunque, por supuesto, no todo el mundo comparte los motivos habitualmente usados para explicarla (Gigerenzer, 1991) y otros consideran que lo que en muchos casos se describe como Desestimación de Probabilidades Previas es en realidad *falacia inversa* (*Inverse Fallacy*) (Hamm, 1993; Koehler, 1996; Villejoubert y Mandel, 2002). En cualquier caso, no es nuestro objetivo entrar en la polémica, solo ejemplificarla.

Ante el siguiente problema (Tversky y Kahneman, 1982):

*“If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming you know nothing about the person’s symptoms or signs?*

*As in the case of confidence and conjunction judgments, subjects were asked for the probability of a single event, that is, that ‘a person found to have a positive result actually has the disease’”.*

Únicamente un 18% de los participantes llega a 0,02, la respuesta normativamente correcta según Tversky y Kahneman, y la mayoría responden 0,95, que se extrae de restarle el 5% de falsos positivos a 100%, quedando un 95% de “verdaderos” positivos.

Cuadro 1. Resolución formal del problema.

Representación gráfica del problema (en azul oscuro la información aportada).

P (Enfermedad) = 1/1000 personas

P (Test + | Enfermedad) = 100% = 1 persona

P (Test + | No Enfermedad) = 999\*5% = 50 personas



Por lo tanto, el tipo de respuesta dada por la mayoría de los sujetos supone desestimar las probabilidades previas (de ahí lo de *desestimación de probabilidades previas*). Además, esta manera “normativamente errónea” de responder resulta extraordinariamente común. Aparece en todo tipo de problemas, condiciones y formatos.

Pero a los partidarios de la *Ecological Rationality* les molesta este enfoque centrado en lo que la mente hace mal. Especialmente porque consideran que la regla de medir usada para determinar qué está bien y qué está mal es errónea. La mayor parte de sus críticas se centran en aspectos teóricos basados en la elección arbitraria de teorías estadísticas normativas y la aplicación incoherente de principios elementales. Por hacer esto algo más concreto, en el ejemplo anterior la crítica más importante se basa principalmente en dos cuestiones. En primer lugar, no se especifica si la persona sobre la que se pregunta fue extraída o no de la población mediante muestreo aleatorio, lo que puede suponer un grave problema, dado que los médicos acostumbran a ver gente de manera no aleatoria (normalmente uno va al médico cuando tiene algún tipo de síntoma), lo que haría que la respuesta mayoritaria (95%) fuera la correcta (Gigerenzer, 1991; Gigerenzer y Murray, 1987, capítulo 5). Dicen, además, que resulta de lo más sencillo acabar con la popular Desestimación de Probabilidades Previas, lo único que hay que hacer es usar frecuencias absolutas y muestreo natural. Kahneman y Tversky lo niegan, sacando a colación situaciones en las que, usando muestreo natural, se aprecia Desestimación de Probabilidades Previas (Camerer, 1990; Grether, 1980, 1992; Griffin y Dukeshire, 1993), y apuntando a los propios resultados de Gigerenzer en los que, según éstos, existen diferencias que él mismo quiso minimizar (Kahneman y Tversky, 1996).

Por otro lado, basándose en un marco teórico estadístico frecuentista, Gigerenzer y Hoffrage, aseguran, además, que los juicios de probabilidad sobre eventos simples, en los que se han basado muchos de los trabajos de Kahneman y Tversky, no tienen sentido, y se apoyan en Finetti:

“*However an individual evaluates the probability of a particular event, no experience can prove him right, or wrong; nor, in general, could any conceivable criterion give any objective sense to the distinction one would like to draw, here, between right and wrong*” (De Finetti, 1989, pag. 174).

La disputa prosigue y los ataques son cada vez más encarnizados. Kahneman y Tversky responden a las críticas desde varios frentes. En primer lugar, destacan la contradicción que supone basarse en el principio del muestreo natural como punto de partida de la superioridad de las frecuencias, cuando en experimentos en los que se ha usado un paradigma de muestreo natural los resultados han contradicho de manera evidente los principios propuestos que harían de las frecuencias naturales el formato elegido. Por ejemplo, contrariamente a la predicción frecuentista, en un experimento en el que la información se mostraba de manera secuencial (emulando lo que sería el muestreo natural), y siendo la frecuencia de una enfermedad común tres veces superior que una rara, los sujetos estimaban la frecuencia de la enfermedad dado el síntoma como si ambas fueran igualmente probables, obviando, por lo tanto, las probabilidades previas (cuya conservación es una de las maravillas de las frecuencias naturales) y basándose en la capacidad diagnóstica del síntoma en su lugar (Gluck y Bower, 1988; Kahneman y Tversky, 1996).

Destacan Kahneman y Tversky que reemplazar juicios de probabilidad por estimaciones de frecuencias y la introducción de muestreo natural no son suficientes para acabar con la Desestimación de Probabilidades Previas (Kahneman y Tversky, 1996), y apuntan a la existencia de otros trabajos en los que este tipo de paradigma ha sido insuficiente para que el rendimiento se ajustara a lo esperado (Estes, Campbell, Hatsopoulos, y Hurwitz, 1989; Nosofsky, Kruschke, y McKinley, 1992). Así que, según ellos, la afirmación que realiza Gigerenzer de que el uso de frecuencias hace que las ilusiones cognitivas desaparezcan no tiene base (Kahneman y Tversky, 1996).

En segundo lugar, se defienden de las acusaciones con respecto al absurdo que supone teóricamente el uso de juicios simples de probabilidad alegando que una parte importante de nuestras decisiones no son más que juicios sobre probabilidades de eventos simples basados en factores intensionales, como creencias (Kahneman y Tversky, 1996). Y es que cosas como someterse a una operación dependen del grado de confianza que uno tiene en que la operación vaya bien, y la decisión no se toma, generalmente, a partir de las frecuencias registradas a partir del muestreo aleatorio de una población de referencia, como exigiría un marco estadístico frecuentista (Kahneman y Tversky, 1996).

La distinción entre factores extensionales e intensionales no se hace explícita en ningún momento en los trabajos sobre cálculo de probabilidad y formatos de representación. En general, los aspectos más cualitativos tienen que ver con la dimensión intensional y los cuantitativos con la extensional. Siguiendo a Johnson-Laird (Johnson-Laird, Legrenzi, Girotto, Legrenzi, y Caverni, 1999), hablamos de cálculo de probabilidad extensional cuando para obtener la probabilidad de un evento nos basamos en las diferentes formas que la situación puede tomar. Hacking (Hacking, 1975), por otro lado, distinguía entre probabilidad aleatoria y epistémica, refiriéndose la primera a frecuencias estables a largo plazo, y la segunda a hasta qué punto alguien cree que algo es cierto (Macdonald, 1986). Esta distinción es paralela a la de probabilidad objetiva-subjetiva y, por otro lado, importante para entender adecuadamente la diferencia entre el cálculo de probabilidad extensional e intensional. El caso del cálculo intensional o no-extensional se podría considerar como una aproximación top-down (desde lo conceptual hacia la realidad) mientras que el extensional sería el método bottom-up (partiendo de la realidad hacia los conceptos).

Es irónico, cuanto menos, que una de las críticas más fuertes que hace Gigerenzer se base en la aplicación estricta de un principio normativo/teórico como es proclamar el sinsentido de los juicios de probabilidad sobre eventos simples, cuando precisamente la adherencia a principios normativos rígidos (y erróneos) es lo que él más le critica a la tradición de los Heurísticos y Sesgos. Especialmente porque, como aseguran Kahneman y Tversky, los juicios (intensionales) sobre eventos simples bien podrían ser una parte, no solo fundamental, sino además inevitable y crucial del día a día.

La polémica llega a niveles de desprecio llamativos (por ambas partes), con acusaciones de uso deliberadamente parcial de problemas y resultados, de plantear posturas como propias del otro bando falsamente y un largo etcétera hasta el punto de distorsionar las posturas teóricas y empíricas. Un ejemplo muy elocuente:

*“The position described by Gigerenzer is indeed easy to refute, but it bears little resemblance to ours. It is useful to remember that the refutation of a caricature can be no more than a caricature of refutation”* (Kahneman y Tversky, 1996, pag. 584).

Con la publicación en 1995 del trabajo de Gigerenzer y Hoffrage, el foco de atención dio un giro radical y pasó a centrarse en dos puntos: la defensa de las frecuencias naturales como formato especialmente apto para el cálculo Bayesiano por la simplificación computacional que suponen, y la existencia de un módulo o algoritmo cognitivo especializado para éstas.

### Cálculo de probabilidad, formato de representación y frecuencias naturales

El día a día está lleno de decisiones que implican algún tipo de inferencia sobre la probabilidad de eventos. Llevamos un paraguas si pensamos que es probable que llueva y buscamos asistencia médica si creemos que nuestra vida corre peligro. Ya que este tipo de decisiones determinan en buena medida nuestro futuro y pueden en ocasiones ser cruciales, lo buenos que somos estimando probabilidad es una pregunta central de la psicología.

Un hecho bien establecido es que el tipo de problemas Bayesianos que se puede ver en el cuadro 2 se resuelve mejor cuando es presentado en forma de frecuencias naturales que como probabilidades de eventos simples (Brase, 2003; Brase y Barbey, 2006; Cosmides y Tooby, 1996; Gigerenzer y Hoffrage, 1995).

Cuadro 2. Ejemplo del material experimental clásico usado por Gigerenzer y Hoffrage, 1995.

**1) Probabilidad Estándar**

La probabilidad de cáncer de mama es de un 1% para mujeres de 40 años que participan en un chequeo rutinario.

Si una mujer tiene cáncer de mama, la probabilidad de que la mamografía dé positivo es de un 80%. Si una mujer no tiene cáncer de mama, la probabilidad de que la mamografía también dé positivo es de un 9,6%.

Una mujer de este grupo de edad da positivo en una mamografía en un chequeo rutinario. ¿Cuál es la probabilidad de que ella tenga realmente cáncer de mama? \_\_\_\_%

**2) Frecuencias Naturales**

Diez de cada 1000 mujeres de 40 años que participan en chequeos rutinarios tienen cáncer de mama.

Ocho de cada 10 mujeres con cáncer de mama darán positivo en una mamografía. De cada 990 mujeres sin cáncer de mama hay también 95 que darán positivo en una mamografía.

Aquí está una nueva muestra representativa de mujeres de 40 años que han dado positivo en una mamografía en chequeos rutinarios. ¿Cuántas de estas mujeres esperas que tengan realmente cáncer de mama? \_\_\_\_ de \_\_\_\_

Ha habido, no obstante, una considerable y extensa discusión sobre los motivos de la aparente superioridad de un formato de representación sobre el otro. En un lado de este debate se sitúan los frecuentistas, defensores de la superioridad de las frecuencias naturales. Pero el frente frecuentista no es uniforme, la *Ecological Rationality* de la que hablamos antes se situaría dentro de un conglomerado de teorías cuyo ámbito común es la defensa de las frecuencias naturales como formato idóneo para el cálculo de probabilidad, al menos Bayesiano. Barbey y Sloman establecen una gradación que nos parece algo más detallada que las posturas que se encuentran en la literatura (Brase, 2007), pero que resulta útil para entender los argumentos teóricos (Barbey y Sloman, 2007). La visión que hemos explicado aquí correspondería al nivel de “Mente como navaja Suiza” (*Mind as Swiss army knife*) (Cosmides y Tooby, 1996; Gigerenzer y Fiedler, 2003; Gigerenzer y Selten, 2001; Todd y Gigerenzer, 2007), el más extremo dentro de las posibilidades frecuentistas. Suavizándose en términos de impenetrabilidad cognitiva, encapsulación de la información o aspiraciones evolucionistas, se encuentran las posturas de “Algoritmo de frecuencias naturales” (*Natural frequency algorithm*) (Gigerenzer y Hoffrage, 1995), “Heurístico de frecuencias naturales*” (Natural frequency heurístic) (Gigerenzer y Hoffrage, 2007; Gigerenzer y Selten, 2001; Tversky y Kahneman, 1974)* y “Heurístico de frecuencias naturales no-evolutivo” (*Non-evolutionary natural frequency heurístic*), respectivamente (Barbey y Sloman, 2007). La única de estas posturas que no incluye los aspectos evolutivos es la última y, hasta donde nosotros sabemos, no es más que una posibilidad teórica sin apoyos significativos.

En cualquier caso, si existiese algo que pudiéramos cualificar como postura frecuentista, y nosotros creemos que sí, ésta, en general, se basaría en dos ejes fundamentales consistentes a lo largo de la práctica totalidad de su literatura:

1) Los algoritmos Bayesianos son más sencillos computacionalmente cuando la información está codificada en forma de frecuencias naturales (Gigerenzer y Hoffrage, 1995).

2) Asumiendo que los humanos han desarrollado a través de la evolución algoritmos cognitivos, éstos habrán sido diseñados para trabajar con frecuencias naturales (Brase, 2003; Brase y Barbey, 2006; Cosmides y Tooby, 1996; Gigerenzer, 1991, 1993; Gigerenzer y Hoffrage, 1995).

Con respecto al primer punto apenas hay discrepancias. Se acepta generalmente que las frecuencias naturales hacen más sencillo el cálculo de probabilidad Bayesiano y la disputa se centra en el porqué. El segundo punto es el más controvertido y, en la forma en la que lo hemos expresado aquí, englobaría lo que entendemos como postura frecuentista general (Brase, 2003; Brase y Barbey, 2006; Cosmides y Tooby, 1996; Gigerenzer y Hoffrage, 1995), en la que se asume que existe algún tipo de estructura (algorítmica, modular o heurística) diseñada por selección natural y que permite computar frecuencias naturales de manera independiente de las habilidades cognitivas generales, con el grado de impenetrabilidad y encapsulamiento que cada cual guste (Barbey y Sloman, 2007). Tal y como han dicho (Chase, Hertwig, y Gigerenzer, 1998, pag. 211): “*Human cognitive algorithms for this type of inference, if they exist, are most likely designed to operate on numerical information in the form in which humans have gathered it over evolutionary history: natural frequencies, that is, absolute frequencies that have not been normalized with respect to the base rates*”.

Los datos a favor de la superioridad de las frecuencias naturales son apabullantes. Ésta se ha demostrado con innumerables problemas (Cosmides y Tooby, 1996; Gigerenzer, 1991; Gigerenzer, 2002; Gigerenzer y Hoffrage, 1995), con niños (Zhu y Gigerenzer, 2006), etc. Pero contrariamente a la predicción principal del campo frecuentista, varios artículos han comunicado situaciones en las que no se encontraba ningún efecto facilitador de las frecuencias sobre las probabilidades (Barbey y Sloman, 2007; Evans, Handley, Perham, Over, y Thompson, 2000; Fiedler, Brinkmann, Betsch, y Wild, 2000; Girotto y Gonzalez, 2001; Johnson-Laird et al., 1999; Macchi, 2000; Macchi y Mosconi, 1998; O’Brien, Roazzi, y Maria da Graça, 2004; Roazzi, O'Brien, y Dias, 2003). Hoffrage y colaboradores (Hoffrage, Gigerenzer, Krauss, y Martignon, 2002) respondieron a las críticas argumentando que algunos de estos autores estaban usando *frecuencias normalizadas* en lugar de *frecuencias naturales* y que en general se usaron estrategias poco elegantes (normalizando frecuencias naturales, etiquetando frecuencias naturales como probabilidades, etc.) para cerrar la brecha entre probabilidades y frecuencias naturales y se renombran principios y ecuaciones ya mostradas por ellos. Estrategias que, por otro lado, surgen de una grave malinterpretación o de una lectura superficial (Hoffrage et al., 2002).

Desde el punto de vista frecuentista, las frecuencias naturales son el resultado del muestreo natural, que es la manera en la que nuestra especie se ha encontrado la información estadística durante milenios, y no es más que a partir de una presentación secuencial de información. Al llegar a la cueva por la noche uno recuerda: hoy vi 14 tigres dientes de sable, de los que 8 eran machos. Las frecuencias naturales, además, no están normalizadas, lo que significa que los datos se aportan en frecuencias absolutas (el número total de observaciones) y no han sufrido una conversión arbitraria (e.g. recodificándolos en base 100). Por ejemplo, en una posible visión alternativa, al llegar a la cueva uno recordaría: hoy vi bastantes tigres dientes de sable, de los que el 57% eran machos.

Por otro lado, según admiten los frecuentistas (Gigerenzer y Hoffrage, 1995), el razonamiento Bayesiano con frecuencias naturales (i.e. absolutas) supone menos pasos de cálculo que con frecuencias relativas o probabilidades (ver también, Howson & Urbach, 1993; Machi & Mosconi, 1998), porque éstas conservan información sobre las probabilidades previas (Hoffrage et al., 2002), factor que también ha sido desestimado por algunos de los críticos.

Sin embargo, algunos estudios también han demostrado que este tipo de simplificación computacional también puede ser obtenida por medios distintos al uso de frecuencias naturales (Evans et al., 2000; Fiedler et al., 2000; Macchi, 2000), aunque puede ser argumentado (tal y como han hecho Hoffrage et al., 2002), que en estos experimentos los formatos probabilísticos imitan el muestreo natural (*natural sampling*) característico de las frecuencias naturales. Desde el punto de vista frecuentista, las frecuencias son la causa de facilitación y no solo una de las posibles maneras de simplificar la computación en problemas Bayesianos.

Si existe un módulo o algoritmo en la mente que ha evolucionado para computar frecuencias naturales y no probabilidades, la facilitación debería ocurrir en cualquier tipo de problema de probabilidades. Se ha demostrado, por ejemplo, que facilitan el razonamiento también en situaciones Bayesianas complejas en las que existen varios predictores, o predictores y criterios con varios valores (Krauss, Martignon, Hoffrage, y Gigerenzer, 2002). También esperaríamos que su efecto fuera, al menos, tan prevalente como el de factores contextuales (e.g. forma exacta de las instrucciones, coincidencia en escala entre instrucciones y respuesta y otros similares). Por poner un ejemplo que aclare a qué nos referimos, en el caso de la coincidencia en escala entre instrucciones y respuesta, si comparamos situaciones en las que hay coincidencia entre el contenido de las instrucciones (el 57% de los tigres dientes de sable son machos) y la respuesta solicitada (57%), con situaciones en las que no (un 0,57 de los tigres dientes de sable son machos), esperaríamos que este factor, que es puramente contextual, fuera menos importante que el formato de representación.

Si no fuera así, y los errores pudieran ser atribuidos a un artefacto en el formato de presentación u otras variables similares, el papel de las frecuencias naturales en particular y del formato de representación en general perdería relevancia. El punto no es trivial, ya que algunos autores se basan en el amplio corpus de resultados relativos a la superioridad de las frecuencias naturales para proclamar que las personas son buenos estadísticos intuitivos (Cosmides y Tooby, 1996).

### Adquisición y representación de información

Según algunos psicólogos evolucionistas, la mente humana ha evolucionado para procesar información en el formato en el que ésta es adquirida en el mundo real (Cosmides y Tooby, 1996; Gigerenzer y Hoffrage, 1995), habiéndose consolidado a lo largo de nuestra historia evolutiva un módulo o algoritmo especializado para el procesamiento de frecuencias naturales (Brase, 2003; Brase y Barbey, 2006; Cosmides y Tooby, 1996; Gigerenzer y Hoffrage, 1995). Ejemplos de una adquisición de información serial-frecuentista se pueden encontrar en ámbitos tan dispares como la caza y recolección animal o las redes neuronales artificiales (Gigerenzer y Hoffrage, 1995).

No obstante, la propuesta frecuentista establece, desde nuestro punto de vista, un paralelismo engañoso entre adquisición de información y representación de la información. No es necesario asumir a priori este paralelismo. Un ejemplo puede ser encontrado en el dominio de las redes neuronales, el mismo que (Gigerenzer y Hoffrage, 1995) citan para ilustrar la adquisición de información secuencial. Aunque es cierto que las redes neuronales adquieren información de manera secuencial, éstas representan la información de manera amodal, distribuida y probabilística (Hinton, McClelland, y Rumelhart, 1986). En un sistema así, habría un coste computacional innecesario asociado a una representación frecuentista de la información.

En cualquier caso, es muy importante tener en cuenta que, a partir del hecho de que la naturaleza secuencial de los formatos frecuentistas puede corresponder con el de la adquisición de información a través del muestreo natural (Gigerenzer y Hoffrage, 1995), no se desprende que la misma correspondencia pueda encontrarse a nivel de representación y manipulación de información. Los defensores de la *Ecological Rationality* basan una parte importante de los argumentos evolutivos a favor de las frecuencias naturales en este tan aventurado como innecesario paralelismo.

### Representación del número

Por otro lado, estudios del desarrollo y transculturales sobre la representación del número en los seres humanos sugieren algunos sistemas mucho menos sofisticados de lo que sería necesario para manejar frecuencias naturales como candidatos a los sistemas evolutivos de representación numérica central (*evolved core number systems*) (Bishop, 1997; S. Dehaene, 1997; S Dehaene, 2001; Feigenson, Dehaene, y Spelke, 2004). Aunque el debate sobre los detalles no está aún cerrado, los sistemas numéricos propuestos (e.g. representación borrosa de números aproximados) son decididamente demasiado simples como para manejar la relativa complejidad de un sistema especializado para las frecuencias naturales tal y como proponen los frecuentistas.

### Complejidad computacional

Comentábamos antes que el primer pilar de la postura frecuentista, el hecho de que el uso de algoritmos Bayesianos sea más sencillo cuando la información está expresada en forma de frecuencias naturales, es apenas controvertido y que la disputa se centra en por qué esto es así y, sobre todo, en el segundo pilar, el evolucionista.

Cuando elaboran sobre la mayor complejidad que supone el cálculo con probabilidades con respecto a las frecuencias naturales, Gigerenzer y Hoffrage (1995) lo hacen desde una postura que parece centrada en distanciar ambos formatos de representación más que en explorar de manera profunda el principio que subyace a las diferencias. Ellos distinguen entre formato de la información y menú de la información. La primera dimensión enfrenta a probabilidades y frecuencias naturales, y la segunda, el menú de la información, se basa en la manera en que la información está segmentada en cada formato y diferencia entre menú estándar y abreviado. Para los problemas del formato estándar de probabilidades se muestran las tres piezas necesarias para la resolución del problema, p(H), p(D|H) y p(D|~H), que son la probabilidad previa – de tener la enfermedad –, la tasa de aciertos (*hit rate*) – probabilidad de dar positivo cuando se tiene la enfermedad – y la tasa de falsas alarmas (*false alarm rate*) – probabilidad de dar positivo cuando no se tiene la enfermedad – respectivamente. En realidad también se necesita p(~H) – probabilidad de no tener la enfermedad –, pero ésta se puede calcular a partir de p(H). En el caso del menú abreviado, que es el que corresponde a frecuencias naturales, solo son necesarias dos piezas de información, d&h (positivo y enfermedad) y d&~h (positivo y no enfermedad) o, según ellos, alternativamente d&h y d (positivos) (Gigerenzer y Hoffrage, 1995, pag. 687).

Cuadro 3. Información en los menús estándar y abreviado para

probabilidades y frecuencias naturales (Gigerenzer y Hoffrage, 1995).

**Menú Estándar**

**Probabilidades**

p(H)

p(D|H)

p(D|~H)

\*p(~H)= 1-p(H)

**Menú Abreviado**

**Frecuencias naturales**

d&h

d&~h

Esta distinción resulta extraña por varios motivos, pero principalmente por la redundancia con respecto a la primera dimensión, el formato de representación. Dicho de otra manera, ellos distinguen entre probabilidades y frecuencias naturales por un lado y entre la información que se aporta en un problema de probabilidades y uno de frecuencias naturales por el otro, aunque eso sí, después cruzan ambos factores creando condiciones de probabilidades con menú abreviado y frecuencias naturales con menú estándar. Encuentran diferencias enormes entre los formatos de representación, pero esto lo hacen comparando condiciones que no solo se diferencian en el formato de la información, sino también en cosas como la clase de referencia (absoluta para frecuencias naturales y relativa para probabilidades), por no hablar de la complejidad computacional, aunque esto último, según ellos, es por definición y, por lo tanto, inevitable. El menú de la información también tiene su influencia, pero ésta es más moderada y por un motivo u otro ha sido ensombrecida históricamente por su hermano mayor, el formato de representación.

### Enredos y desenredos

La discusión sobre formatos de representación es enormemente compleja y, a pesar de que hay bastante acuerdo sobre el rol de algunos factores como la desambiguación semántica (Hertwig y Gigerenzer, 1999) y la clarificación de la clase de referencia (Gigerenzer et al., 1991), las confusiones y malinterpretaciones por parte del frente opositor a los frecuentistas, algunas críticas sin base empírica por parte de los frecuentistas (e.g. influencia de la muestra total), etc., han hecho prácticamente imposible un acuerdo (Barbey y Sloman, 2007; Brase, 2007; Brase y Barbey, 2006; Gigerenzer y Hoffrage, 1995; Girotto y Gonzalez, 2001, 2002; Hoffrage et al., 2002; Johnson-Laird et al., 1999). Desde nuestro punto de vista, los temas críticos que hacen imposible dirimir el asunto son, la mala conceptualización de la complejidad computacional, por un lado, y la confusión entre complejidad computacional y el estatus evolutivamente especial de las frecuencias, por otro lado. La operacionalización de complejidad computacional que realizan los frecuentistas es poco útil, porque existe un solapamiento entre los pasos de cálculo y los menús de la información (para la condición de frecuencias naturales, el número de pasos de cálculo es idéntico para los menús abreviado y estándar).

Además, un inconveniente que surge al usar problemas Bayesianos para demostrar las ventajas de las frecuencias naturales sobre otros formatos de representación es el provocado por las dificultades que se encuentran a la hora de desenredar la influencia de la complejidad matemática de la del formato de representación, ya que, como hemos comentado antes, la resolución de problemas Bayesianos (del tipo empleado habitualmente) con frecuencias naturales es computacionalmente más sencilla que con probabilidades. En el cuadro siguiente se pueden ver los algoritmos Bayesianos necesarios para resolver uno y otro problema y cómo la dificultad difiere entre ellos.

Cuadro 4: Algoritmos Bayesianos para la resolución de problemas de probabilidades y frecuencias naturales.

**Probabilidades**



**Frecuencias naturales**



Es difícil, por lo tanto, diferenciar entre los dos ejes propuestos por los frecuentistas y determinar si las diferencias que se encuentran son causadas por la mayor simplicidad de las frecuencias naturales al resolver problemas Bayesianos o por su presunto especial estatus evolutivo, el cual no se puede probar. Además, es complicado hablar sobre la relevancia del otro factor manipulado por ellos, el menú de la información, cuando existe un solapamiento con la complejidad aritmética.

En este trabajo pretendemos deshacer este enredo. Nuestra intención es comprobar empíricamente algunas de las afirmaciones que se han realizado en el campo y, sobre todo, buscar las causas subyacentes a las diferencias demostradas entre los distintos formatos de representación. Es una tarea que creemos factible, pero no debemos olvidar la cantidad ingente de publicaciones y el profundo desacuerdo que aún hoy, después de más de una década, reina sobre las tierras del formato de representación y cálculo de probabilidad. Idealmente nuestros resultados deberían aportar un nuevo marco interpretativo, o mejor, un marco más claro que contribuya a cerrar la polémica, evitando convertirse en otro conjunto de resultados negando o apoyando la superioridad de uno u otro formato de representación para este o aquel tipo de cálculo de probabilidad.

Más concretamente, intentaremos comprobar si algo como la complejidad computacional o algorítmica definida como los pasos de cálculo puede ser suficiente para explicar la aparente superioridad de las frecuencias naturales, más allá de menús de la información, módulos especializados y formatos de representación.

### Probabilidades simples y Bayesianas

La facilitación de la que hablan los frecuentistas no se da únicamente en los problemas clásicos usados por Gigerenzer y Hoffrage. Asimismo, como veíamos más arriba, se ha demostrado que facilitan el razonamiento también en situaciones Bayesianas complejas en las que existen varios predictores, o predictores y criterios con varios valores (Krauss et al., 2002).

Parte de los argumentos de los partidarios de la *Ecological Rationality* en defensa de las frecuencias naturales como formato seleccionado por la evolución tienen su base en argumentos de sentido común sobre la ubicuidad de las frecuencias a lo largo de nuestra historia evolutiva. Si tenemos en consideración que para realizar un cálculo posterior o Bayesiano son necesarias como mínimo 2 probabilidades simples (o al menos más simples), y que es una verdad analítica que la presencia de las probabilidades simples es mayor que la de posteriores, a priori no se debería descartar que el presunto módulo o algoritmo cognitivo existente para trabajar con frecuencias naturales facilite también el cálculo con probabilidades simples, o que exista un segundo módulo o algoritmo para éstas.

### Objetivos e Hipótesis

Dividiremos la exposición de los resultados en dos grandes bloques que tratarán de manera diferenciada, por un lado, el cálculo de probabilidades simples, y, por otro lado, el cálculo de probabilidades Bayesianas usando versiones ligeramente modificadas de las tareas clásicas. Esta separación obedece a la voluntad de que, en primer lugar, los resultados sean extrapolables más allá de las tareas de cálculo de probabilidad Bayesiano y, en segundo lugar, a la necesidad de simplificar conceptualmente el estudio del papel de los formatos de representación en el cálculo de probabilidad. Miraremos también a través de esta separación, veremos cómo el principio fundamental que regula los resultados de los experimentos de la primera parte, aquella referida a probabilidades simples, se aplica de manera elegante a los datos de la segunda parte, en la que se trata el cálculo Bayesiano.

Nuestros objetivos son muy claros. Pretendemos despojar de ambigüedades, complejidad y variables extrañas el estudio del cálculo de probabilidad, particularmente el Bayesiano. Han sido tantos los pasos en falso, malinterpretaciones, afirmaciones sin base empírica, que creemos que es necesario comprobar, a partir de principios lo más sencillos posibles, dónde nos encontramos realmente. Por ejemplo, hasta el momento, las afirmaciones sobre la superioridad de las frecuencias naturales, especialmente desde la *Ecological Rationality*, han resultado casi indistinguibles de los argumentos evolutivos. ¿Son las frecuencias naturales más sencillas porque la evolución ha creado un módulo o algoritmo especializado para su tratamiento por la omnipresencia de éstas? O. por el contrario, ¿su mayor simplicidad intrínseca ha provocado que la selección natural las seleccionara? ¿O tal vez los argumentos evolutivos son un añadido innecesario?

Las hipótesis de fondo que nos mueven se podrían resumir en los siguientes puntos:

1. Las diferencias entre probabilidades y frecuencias naturales se dan por la disparidad en complejidad de ambos formatos en los problemas Bayesianos. Al usar probabilidades simples estas diferencias deberían desaparecer.
2. Cualquier factor que añada complejidad a la resolución de un problema (sea éste con probabilidades simples o posteriores) afectará al rendimiento de los sujetos independientemente del formato de representación.
3. Si no existiera un módulo o algoritmo especializado para las frecuencias naturales, las habilidades cognitivas generales deberían influir en la correcta resolución de los problemas Bayesianos.
4. La complejidad computacional o aritmética, entendida como pasos de cálculo necesarios para resolver el problema, podría ser una herramienta útil para predecir el comportamiento de los sujetos y explicar diferencias basadas en formatos de representación, clases de referencia, estructura del problema, etc.

# EXPERIMENTOS

## 1. Factores contextuales y formatos de representación.

En esta sección ponemos a prueba la hipótesis frecuentista usando probabilidades simples en lugar de condicionadas. La aparición de las probabilidades simples en este primer bloque de los resultados responde a que, ya que para los proponentes de la *Ecological Rationality* a lo largo de nuestra historia evolutiva se ha desarrollado un algoritmo cognitivo especializado en el tratamiento de frecuencias naturales, éstas deberían también mejorar el rendimiento en los problemas de probabilidad simple, donde no es necesaria simplificación computacional. Este extremo, que nosotros sepamos, no ha sido comprobado hasta ahora, a pesar de su importancia, ya que, como decíamos, cada vez que uno de nuestros antepasados se encontró con las frecuencias necesarias para calcular una probabilidad Bayesiana, previamente debió encontrar, al menos, dos frecuencias simples.

La probabilidades simples (e.g. la simple relación entre casos favorables y posibles) han sido históricamente descartadas, porque son fáciles de calcular y generalmente no tan dadas al error (e.g. probabilidad de que en el siguiente lanzamiento de una moneda salga cara). Sin embargo, tienen una ventaja importante y es que nos permiten controlar la confusión creada por la complejidad computacional, en el sentido de que en probabilidades simples no hay diferencia en términos de complejidad entre problemas en formato probabilista o frecuentista. Además, un paradigma más sencillo debería ayudar a minimizar otros factores como la desambiguación semántica y la clarificación de la clase de referencia. Por último, si las frecuencias facilitan el razonamiento lo deberían hacer tanto en problemas Bayesianos como de probabilidad simple. Es poco probable que nuestros ancestros desarrollaran dos mecanismos separados e independientes para probabilidades simples y condicionadas.

Un tipo de problema de probabilidad simple en el que se han encontrado tasas de error considerables es la inferencia a partir de “rachas”. En estas tareas, los participantes son informados de una secuencia de eventos y se les pide que predigan el siguiente evento de la serie o que calculen la probabilidad de eventos subsiguientes (e.g. tiradas de moneda). Este procedimiento tiene, además, la ventaja de que puede ser considerado como un buen ejemplo de muestreo natural.

En este primer bloque esperamos dilucidar si la teórica ventaja de las frecuencias naturales se da también usando problemas simples, y qué determina exactamente las diferencias, si las hubiera. ¿Son relevantes cosas como la presencia o ausencia de la clase complementaria o el emparejamiento entre instrucciones y respuesta más allá del formato de representación?

### Probabilidades de eventos simples I

El primer experimento pone en juego juicios simples de probabilidad sobre series de eventos en problemas referidos a la distribución de pimientos picantes y no picantes en platos con diez pimientos de Padrón (de los que decíamos que un 10% eran picantes). Elegimos pimientos de Padrón para estos experimentos porque, como es bien sabido, estos pimientos no tienen ningún signo externo que dé pistas sobre si picarán o no. A su vez, éstos crean de forma espontánea un muestreo aleatorio. Solo un cocinero omnisciente podría usar un procedimiento de selección representativa, cogiendo un pimiento picante y 9 no picantes para cada plato. Los sujetos debían calcular la probabilidad de que el primer y último pimiento del plato fueran picantes (en este último caso sin que hubiera picado ninguno de los otros 9 pimientos). El resultado crítico era la probabilidad de que el último pimiento fuera picante.

En resumidas cuentas, en este primer experimento comparamos frecuencias naturales con probabilidades en una tarea de juicios de probabilidades simples sobre series de eventos. Por lo que hemos comentado con anterioridad, a partir de los principios de la teoría frecuentista se podría deducir lógicamente que los resultados serán superiores en esta tarea cuando es presentada en forma de frecuencias naturales (ya que, como decíamos, es analíticamente cierto que nuestros antepasados vieron más frecuencias simples que posteriores).

#### Método

#### Participantes y diseño

Un total de 80 estudiantes de psicología de la Universidad de Málaga completaron un cuestionario en clase. Fueron asignados de forma aleatoria a dos grupos correspondientes al factor entre-sujetos, *Formato de representación* (frecuencias/probabilidades). Todos recibieron preguntas sobre el primer y último pimiento de un plato con 10 pimientos de Padrón. Esto definió el factor intra-sujetos *Ensayo* (primer/último pimiento).

Tabla 1. Variables independientes usadas y representación esquemática del material.

**Variable independiente**

Entre-sujetos:

* Formato de representación
  + Probabilidades (0,1 – 0,9)
  + Frecuencias (1 de 10 – 9 de 10)

Intra-sujetos:

* Ensayo (primer/último pimiento)

#### Procedimiento y materiales

Los sujetos respondían a dos preguntas: la primera, sobre la probabilidad de que el primer pimiento cogido de un plato lleno picara, y la segunda, sobre la probabilidad de que el último pimiento del plato picara (habiendo sido los otros 9 no picantes). Las probabilidades se establecieron como 0,1 (1 de cada 10) para los picantes y 0,9 (9 de cada 10) para los no picantes. Por lo tanto, 0,1 en cualquiera de sus variantes era siempre la respuesta correcta. En la tabla siguiente se puede ver el material experimental:

Tabla 2. Material del experimento 1 (en cursiva).

*Te vamos a presentar un problema de razonamiento. El objetivo es estudiar la forma de razonar de las personas. Hay dos preguntas. Responde a la primera antes de pasar a la segunda y no vuelvas atrás.*

*Según un estudio de la Xunta de Galicia, en la producción de este año de pimientos de Padrón (normalmente algunos de estos pimientos pican y otros no),*

* **Probabilidades**

*la probabilidad de que piquen es del 0,1 y de que no piquen, del 0,9.*

*En un bar te sirven un plato con diez de estos pimientos (obtenidos de la misma cosecha a que se refiere el estudio).*

*Pregunta 1: Imagínate que coges un pimiento cuando el plato está lleno.*

*¿Cuál es la probabilidad (en porcentaje) de que sea picante?*

*\_\_\_\_ %*

*Pregunta 2: Imagínate que coges el último pimiento del plato cuando los otros nueve no picaban.*

*¿Cuál es la probabilidad (en porcentaje) de que sea picante?*

*\_\_\_\_ %*

* **Frecuencias**

*de cada diez, 1 es picante y 9 no.*

*En un bar te sirven un plato con diez de estos pimientos (obtenidos de la misma cosecha a que se refiere el estudio).*

*Pregunta 1: Imagínate que coges un pimiento cuando el plato está lleno.  
¿Qué posibilidades tienes de que sea picante?*

*\_\_\_\_ de \_\_\_\_*

*Pregunta 2: Imagínate que coges el último pimiento del plato cuando los otros nueve  
no picaban.*

*¿Qué posibilidades tienes de que sea picante?*

*\_\_\_\_ de \_\_\_\_*

*GRACIAS POR TU COLABORACIÓN*

La comparación entre el primer y último pimiento se realiza con la única finalidad de comprobar la validez de nuestro paradigma para detectar diferencias.

#### Resultados

Los porcentajes de respuestas correctas se presentan en la tabla 3. El análisis de varianza (ANOVA) con medidas repetidas para el factor *Formato de representación* (frecuencias/probabilidades) y el factor intra-sujetos *Ensayo* (primer/último pimiento) mostró un efecto principal significativo de *Ensayo* (F(1, 78)= 68,65, MSE= 0,153, p<0,0001, η2=0,47). Los participantes respondieron mejor a la pregunta sobre el primer pimiento (86% aciertos) que a la pregunta sobre el último pimiento (35% aciertos), validando el uso de este paradigma de probabilidades sencillas. También se encontró un efecto principal de *Formato de representación* (F(1, 78)= 10,14, MSE= 0,18, p<0,01, η2=0,12) dado que el rendimiento de los participantes era mayor con probabilidades (71% aciertos) que con frecuencias (50% aciertos). Centrándonos en el resultado crítico (probabilidad de que el último pimiento sea picante), confirmamos un efecto significativo de *Formato de representación* (F(1, 78)= 5,75, MSE= 1,25, p<0,02, η2=0,07), donde los sujetos respondían de forma más precisa con probabilidades (47% aciertos) que con frecuencias (22% aciertos).

En la tabla siguiente se ve con claridad cómo cuando los participantes tenían más dificultades erraban atribuyendo una probabilidad mayor en sus respuestas. Contrariamente a lo que predice la hipótesis frecuentista, las frecuencias naturales incrementaron considerablemente la tasa de error sobre las probabilidades en esta tarea de cálculo de probabilidad simple.

Tabla 3. Porcentaje de respuestas correctas y (entre paréntesis) la media de la probabilidad estimada por los sujetos para cada condición experimental en el experimento 1.

**Primer pimiento Último pimiento N**

**Frecuencias**  77 (17) **22** (71) 40

**Probabilidades**  95 (13) **47** (55) 40

Imaginemos que se creara una representación analógica de la situación, tal y como proponen algunas teorías (ver por ejemplo, Johnson-Laird et al., 1999). En la condición de frecuencias naturales, los sujetos se representarán una sola posibilidad, 1 pimiento picante y 9 no picantes. En realidad, esto no es más que una representación de la cosecha de pimientos en una cierta escala. Sin embargo, los modelos mentales se actualizan con la lectura del discurso de manera que cuando los sujetos tienen que responder a la pregunta sobre el ultimo pimiento (habiendo sido los otros 9 no picantes), la coincidencia en escala les llevará a emparejar su representación de la cosecha de pimientos con la descripción del plato, eliminando simplemente los 9 pimientos no picantes, pues ya no están en el plato. Dicho de otra forma, si los primeros 9 pimientos no eran picantes, sólo queda 1, que es el que nos hemos representado como picante. Podemos, por lo tanto, explicar el mayor rendimiento en la condición de probabilidades a partir de los errores derivados de la *falacia del jugador* (*gamblers’ fallacy)*, que se producen al existir emparejamiento entre las instrucciones y las posibilidades representadas. La *falacia del jugador* es un falacia lógica por la que se cree erróneamente que los sucesos pasados afectan a los futuros en cuestiones regidas por el azar – por ejemplo, al lanzar monedas al aire, pensar que después de varias caras es más probable que salga una cruz – (este fenómeno fue descrito por primera vez por Laplace en 1814; Laplace, 1995).

### 1.2. Probabilidades de eventos simples II

En el experimento 1 encontramos que los participantes eran mejores con probabilidades que con frecuencias naturales. Este resultado es ciertamente un reto para la aproximación frecuentista, pero demasiado preliminar como para poder llegar a alguna conclusión. Se podría argumentar que nuestras probabilidades comparten ciertas similitudes con las frecuencias naturales, ya que las dos probabilidades dadas (0,1 y 0,9) pueden ser sumadas computando la muestra total prácticamente sin esfuerzo, lo que facilitaría la computación, tal y como (Hoffrage et al., 2002, p. 350) dijo, “*Providing the total sample (...) serves as a starting point to mimic the procedure of natural sampling, thereby facilitating computational demands considerably”*.Por otro lado, aunque esto podría explicar la igualdad entre ambos resultados no habla sobre la superioridad que muestran las probabilidades en el primer experimento.

Así pues, con la finalidad de evitar las críticas concernientes a las similitudes entre el formato de nuestras probabilidades y las frecuencias naturales, en este experimento la condición de probabilidades mimetiza lo que Gigerenzer y colaboradores llamaron formato de probabilidades estándar (*standard probability format)* (Gigerenzer y Hoffrage, 1995), siendo expresado como probabilidades de eventos simples (*“el 10% de ellos eran picantes”*).

#### Método

#### Participantes y diseño

Un nuevo grupo de 60 estudiantes de la misma población del primer experimento participaron bajo las mismas circunstancias (contestando un cuestionario en clase). Fueron asignados de forma aleatoria a dos grupos. El primer grupo recibió el problema en forma de frecuencias naturales y el segundo como un porcentaje simple.

El diseño del experimento era de nuevo un 2x2, (frecuencias/probabilidades) x (primer/último pimiento).

Tabla 4. Variables independientes usadas y representación esquemática del material.

**Variable independiente**

Entre-sujetos:

* Formato de representación
  + Probabilidades (10%)
  + Frecuencias (1 de 10 – 9 de 10)

Intra-sujetos:

* Ensayo (primer/último pimiento)

#### Procedimiento y materiales

El procedimiento a seguir en este segundo experimento fue idéntico al del primero. El único cambio se dio en los materiales, y es que en la condición de probabilidades dejamos de presentar explícitamente la probabilidad de que un pimiento no fuera picante (la categoría complementaria). Para evitar dar facilidades a la hora de computar la muestra total mostrábamos únicamente la probabilidad de que el pimiento fuera picante. Como ya hemos dicho, hacíamos esto con un sólo porcentaje (*“el 10% de ellos eran picantes”*).

Más concretamente, en las instrucciones de la condición de probabilidades se cambió [“*la probabilidad de que piquen es del 0,1 y de que no piquen, del 0,9”*] por [“*la probabilidad de que piquen es del 10%”*]. Los formatos de respuesta y todo lo demás se mantuvo igual.

#### Resultados

Los resultados de este experimento se muestran en la tabla 5. El ANOVA mostró una interacción significativa de *Ensayo* y *Formato de representación* (F(1, 58)= 11,87, MSE= 0,12, p< 0,002, η2=0,17), y un efecto principal de *Ensayo* (F(1, 58)= 43,89, MSE= 0,12, p< 0,0001, η2=0,43). Los participantes respondieron mejor a la pregunta sobre el primer pimiento (80% aciertos) que a la pregunta sobre el último pimiento (38% aciertos). La interacción se daba principalmente por el mayor acierto de los participantes en la pregunta crítica, probabilidad de que el último pimiento fuera picante, en la condición de probabilidades (57% aciertos), frente a la condición de frecuencias naturales (20% aciertos), F(1, 58)= 9,61, MSE= 0,21, p< 0,003, η2=0,12).

Tabla 5. Porcentaje de respuestas correctas y (entre paréntesis) la media de la probabilidad estimada por los sujetos para cada condición en el experimento 2.

**Primer pimiento Último pimiento N**

**Frecuencias**  83 (15) **20** (68) 30

**Probabilidades**  77 (10) **57** (30) 30

Los resultados hasta el momento confirman los del experimento 1. En este tipo de problemas de cálculo de probabilidad simple los formatos porcentuales son mejores sistemáticamente que los formatos frecuentistas. Sin embargo, hay algunas críticas que se podrían hacer en lo que respecta a sutiles diferencias contextuales entre los distintos formatos y sobre la interpretación probabilística de nuestras instrucciones. Las veremos en el siguiente experimento.

### 1.3. Probabilidades simples y factores contextuales

En los experimentos anteriores se utilizaba un formato de respuesta que pedía a los participantes que hicieran una predicción sobre el siguiente evento de una serie. Ésta es de hecho una tarea ecológica y, probablemente, la predicción de los “eventos siguientes” es el propósito fundamental de las habilidades estadísticas de la mente humana (Kahneman y Tversky, 1996). De todos modos, se puede argumentar que este énfasis en la validez ecológica puede haber distorsionado nuestro formato de respuesta frecuentista en el sentido en el que las frecuencias se deberían referir a sets de eventos y no a eventos simples. El experimento 3 fue diseñado para superar esta dificultad y poder realizar una comparación más justa entre probabilidades y frecuencias.

Además, pese a que, como hemos explicado, es imposible no seleccionar una muestra aleatoria de pimientos de Padrón (ya que éstos no tienen ningún signo externo que indique si pican o no), hicimos explícito en las instrucciones que la muestra de pimientos del plato fue seleccionada aleatoriamente a partir de la población descrita.

Finalmente, realizamos varias manipulaciones para comprobar empíricamente la relevancia de algunos factores contextuales que pudieron influir en los resultados previos. Añadimos dos dimensiones a los problemas anteriores, Categoría Complementaria Explícita vs. Implícita (i.e. *CC Explícita vs. Implícita*) y Emparejamiento entre Instrucciones y Respuesta (i.e. *Emparejamiento I-R*).

En la primera, *CC Explícita vs. Implícita*, manipulamos la presencia o ausencia de la información sobre los pimientos no picantes, la categoría complementaria a los pimientos picantes. En un caso quedaba una probabilidad normalizada pura referida a eventos simples [“*la probabilidad de que piquen es del 0,1”*] y en el otro, una probabilidad presuntamente cuestionable por la facilidad de computar la muestra total y, por lo tanto, por su similitud con las frecuencias naturales [“*la probabilidad de que piquen es del 0,1 y de que no piquen, del 0,9”*]. Los proponentes de las teorías frecuentistas y de la llamada *Ecological Rationality* han usado esto como crítica (Hoffrage et al., 2002), pero hasta donde nosotros sabemos, no hay prueba empírica de su influencia.

En la segunda dimensión, *Emparejamiento I-R*, controlamos la existencia o ausencia de correspondencia (en escala) entre la probabilidad dada en las instrucciones y la respuesta correcta. Siendo la respuesta correcta 10%, comparábamos la condición con correspondencia [“*la probabilidad de que piquen es del 10%”*] con la de no correspondencia [“*la probabilidad de que piquen es del 0,1”*]. En este último caso, los sujetos tenían que realizar un cambio de escala, de 0,1 a 10%, para dar la respuesta correcta, ya que se solicitaba un porcentaje.

#### Método

#### Participantes y diseño

Un nuevo grupo de 313 estudiantes de psicología de la Universidad de La Laguna fueron asignados aleatoriamente a ocho grupos distintos. Las preguntas sobre el primer y último pimiento siguieron definiendo el factor intra-sujetos *Ensayo* (primer/último pimiento), y los factores entre-sujetos fueron: *Formato de representación* (frecuencias/ probabilidades), *CC Explícita vs. Implícita* (explícita/implícita) y *Emparejamiento I-R* (emparejamiento/no emparejamiento).

Tabla 6. Variables independientes usadas y representación esquemática del material.

**Variable independiente**

Entre-sujetos:

* Formato de representación
  + Probabilidades (10% ó 0,1)
  + Frecuencias (10 de cada 100 ó 1 de cada 10)
* Categoría complementaria
  + Explícita (10% picantes y 90% no picantes)
  + Implícita (10% picantes)
* Emparejamiento entre Instrucciones y Respuesta (I-R)
  + Emparejamiento I-R (10%)
  + No Emparejamiento I-R (0,1)

Intra-sujetos:

* Ensayo (primer/último pimiento)

#### Procedimiento y materiales

De nuevo el procedimiento fue idéntico al de los experimentos anteriores. Fueron, sin embargo, introducidos dos cambios sobre los materiales del experimento 2. En primer lugar, como ya hemos comentado, hicimos explícito que los pimientos eran seleccionados aleatoriamente, cambiando la frase de las instrucciones de: [“*(obtenidos de la misma cosecha a que se refiere el estudio).”*]a [“*(obtenidos al azar de la misma cosecha a que se refiere el estudio).”*]. En segundo lugar, modificamos el formato de respuesta para evitar distorsiones en la interpretación de las frecuencias dejando claro que la respuesta tenía que basarse en un conjunto de eventos y no en un evento simple.

En la tabla siguiente se puede ver el material experimental:

Tabla 7. Material del experimento 3 (en cursiva).

*Te vamos a presentar un problema de razonamiento. El objetivo es estudiar la forma de razonar de las personas. Hay dos preguntas. Responde a la primera antes de pasar a la segunda y no vuelvas atrás.*

*Según un estudio de la Xunta de Galicia, en la producción de este año de pimientos de Padrón (normalmente algunos de estos pimientos pican y otros no),*

* Probabilidades
  + Explícita
    - Emparejamiento I-R

*[la probabilidad de que piquen es del 10% y de que no piquen, del 90%]*

* + - No Emparejamiento I-R

*[la probabilidad de que piquen es de 0,1 y de que no piquen, de 0,9]*

* + Implícita
    - Emparejamiento I-R

*[la probabilidad de que piquen es del 10%]*

* + - No Emparejamiento I-R

*[la probabilidad de que piquen es de 0,1]*

[COMÚN Probabilidades]

*En un bar te sirven un plato con diez de estos pimientos (obtenidos al azar de la misma cosecha a que se refiere el estudio).*

*Pregunta 1: Imagínate que coges un pimiento cuando el plato está lleno.*

*Si hicieras eso 100 veces, ¿Cuál sería la probabilidad (en porcentaje) de que fuera picante?*

\_\_\_\_ %

*Pregunta 2: Imagínate que coges el último pimiento del plato cuando los otros nueve  
no picaban.*

*Si hicieras eso 100 veces, ¿Cuál sería la probabilidad (en porcentaje) de que fuera picante?*

\_\_\_\_ %

[/COMUN Probabilidades]

* Frecuencias
  + Explícita
    - Emparejamiento I-R

*[de cada cien pimientos, 10 son picantes y 90 no]*

* + - No Emparejamiento I-R

*[de cada diez pimientos, 1 es picante y 9 no]*

* + Implícita
    - Emparejamiento I-R

*[de cada cien pimientos, 10 son picantes]*

* + - No Emparejamiento I-R

*[de cada diez pimientos, 1 es picante]*

[COMUN Frecuencias]

*En un bar te sirven un plato con diez de estos pimientos (obtenidos al azar de la misma cosecha a que se refiere el estudio).*

*Pregunta 1: Imagínate que coges un pimiento cuando el plato está lleno.*

*Si hicieras eso 100 veces, con qué frecuencia cogerías pimientos picantes.*

*Número de veces\_\_\_\_\_\_\_*

*Pregunta 2: Imagínate que coges el último pimiento del plato cuando los otros nueve  
no picaban.*

*Si hicieras eso 100 veces, con qué frecuencia cogerías pimientos picantes.*

*Número de veces\_\_\_\_\_\_\_*

[/COMUN Frecuencias]

#### Resultados

Los resultados se presentan en la tabla 8. En este experimento el ANOVA también mostró un efecto principal de *Ensayo* (F(1, 305)= 67,1, MSE= 0,14, p< 0,0001, η2= 0,18), respondiendo mejor los participantes a la pregunta sobre el primer pimiento (79% aciertos) que a la pregunta sobre el último pimiento (54% aciertos). También encontramos una interacción significativa entre *Ensayo* y *CC Explícita vs. Implícita* (F(1, 305)= 4,84, MSE= 0,14, p< 0,029, η2= 0,016), y entre *Ensayo* y *Emparejamiento I-R* (F(1, 305)= 5,18, MSE= 0,14, p< 0,024, η2= 0,017).

La primera interacción fue creada por dos tendencias sutiles e individualmente no significativas. En la pregunta sobre el primer pimiento, la presencia explícita de la *Categoría Complementaria* producía una cierta ventaja (83% aciertos frente a 75% aciertos; t’(311)= -1,623, p= 0,106) (de aquí en adelante informamos de t’ cuando no existe homogeneidad de varianzas), mientras que esto resultaba desventajoso en la pregunta sobre el último pimiento (51% aciertos versus 56% aciertos) t(311)= 0,869, p= 0,385. Este mismo patrón se puede observar comparando la condición de probabilidades en los experimentos 1 (CC) y 2 (no CC). Parece que la presencia de la *Categoría Complementaria* contribuía a un ligero incremento de la *falacia del jugador* (aquella falacia lógica por la que se cree erróneamente que los sucesos pasados afectan a los futuros en cuestiones regidas por el azar) al responder a la pregunta sobre el último pimiento habiendo sido los otros nueve no picantes.

La segunda interacción fue producida por las respuestas significativamente mejores de los participantes a la respuesta crítica, “probabilidad de que el último pimiento fuera picante”, en la condición de *Emparejamiento I-R* (60% aciertos, Intervalo de confianza= 52,6/68,9) que en la de *no Emparejamiento I-R* (48% aciertos, Intervalo de confianza= 40,6/55,8; F(1, 305)= 4,71, MSE= 0,25, p< 0,04, η2= 0,015).Siendo la respuesta correcta 10 de cada 100, la condición de *Emparejamiento I-R* se daría cuando las instrucciones fueran [“de cada 100 pimientos, 10 son picantes”] y la de *no Emparejamiento I-R* cuando fueran[“de cada 10 pimientos, 1 es picante”].

En lo que respecta al formato de representación, no había diferencias entre la condición de probabilidades (50% aciertos, Intervalo de confianza= 42,5/58,2) y la de frecuencias naturales (58% aciertos, Intervalo de confianza= 50,7/66,4), F(1,305) = 2,08, MSE= 0,25, p= 0,150, Potencia= 0,302). Tampoco había diferencias entre *CC Explícita* (52% aciertos, Intervalo de confianza=43,7/60) y *CC. Implícita* (57% aciertos, Intervalo de confianza=49,5/64,7), F(1,305) = 0,843, MSE= 0,2, p= 0,36, Potencia= 0,150).

Los resultados de este experimento cuestionan la ventaja de las frecuencias sobre las probabilidades o viceversa. El foco se centra, sin embargo, en la relativa saliencia de factores contextuales como el emparejamiento entre el formato de las instrucciones dadas y la respuesta requerida (*Emparejamiento I-R*).

La manipulación sistemática de *CC Explícita vs. Implícita* (refiriéndose a situaciones donde los dos términos son mostrados [“*la probabilidad de que piquen es del 10% y de que no piquen, del 90%*”] o bien sólo el primer termino [“*la probabilidad de que piquen es del 10%*”] respectivamente) no muestra ningún efecto significativo, a pesar de la presunta similitud del primer caso con frecuencias naturales (y la teórica simplificación computacional que ello conlleva). Decíamos antes que los proponentes de las teorías frecuentistas y de la llamada *Ecological Rationality* han criticado el uso de probabilidades que mostraban de manera explícita la categoría complementaria (Hoffrage et al., 2002) sobre la base de que esta situación ayuda a equiparar éstas con frecuencias naturales. Nuestros resultados, desde luego, no apoyan este extremo con probabilidades simples.

Tabla 8. Porcentaje de respuestas correctas y (entre paréntesis) la media de la probabilidad estimada por los sujetos para cada condición experimental en el experimento 3.

**Primer pimiento Último pimiento N**

**Frecuencias**

**CC Explícita**

**Emparejamiento I-R**  88 (14) 62 (44) 34

**no Emparejamiento I-R**  74 (21) 44 (29) 43

**CC Implícita**

**Emparejamiento I-R**  71 (14) 71 (21) 35

**no Emparejamiento I-R**  80 (16) 57 (34) 44

**Probabilidades**

**CC Explícita**

**Emparejamiento I-R**  82 (15) 53 (40) 34

**no Emparejamiento I-R**  89 (13) 49 (41) 35

**CC Implícita**

**Emparejamiento I-R**  73 (26) 57 (37) 44

**no Emparejamiento I-R**  77 (11) 43 (34) 44

En resumen, un factor contextual como el emparejamiento entre el formato de las instrucciones dadas y la respuesta requerida (*Emparejamiento I-R*) se muestra como relevante en una tarea de cálculo de probabilidades simples mientras que el *Formato de representación* (frecuencias/probabilidades) no.

### Discusión: Factores contextuales y formatos de representación con probabilidades simples.

En estos experimentos se abren algunas brechas en las principales predicciones frecuentistas. No solo los participantes no rendían más con frecuencias naturales que con probabilidades, sino que, de hecho, factores contextuales como la correspondencia (en escala) entre instrucciones y respuesta eran determinantes de cara a su rendimiento. Se podrían hacer, no obstante, algunas alegaciones ad hoc a favor de la aproximación frecuentista.

En primer lugar, se podría argumentar que los formatos frecuentistas facilitan el razonamiento Bayesiano y no el razonamiento probabilístico simple. Esto es, en principio, consistente con la propuesta de Gigerenzer y Hoffrage (1995) de que la complejidad computacional es responsable del efecto de facilitación, pero no con la afirmación de Cosmides y Tooby (1996) sobre la incapacidad intrínseca de la mente humana para lidiar con probabilidades a no ser que éstas se presenten en forma de frecuencias.

Un problema con la hipótesis de la complejidad computacional es la confusión a la que nos referimos antes. Si la única proposición comprobable del campo frecuentista es que cuando los problemas Bayesianos son expuestos en un formato que simplifica la computación (con frecuencias naturales, “*number of chances”*, etc.) los participantes resuelven mejor los problemas, no hay mucho que decir al respecto. Que la multiplicación es más sencilla con números arábigos que con números romanos no es algo nuevo en el siglo XXI (véase por ejemplo (Hoffrage et al., 2002). Pero mientras que la complejidad computacional parece ser la única proposición falsable de su propuesta, los artículos frecuentistas están repletos con otras afirmaciones relativas a la evolución cognitiva. Por ejemplo: “*The evolutionary argument that cognitive algorithms were designed for frequency information, acquired through natural sampling, has implications for the computations an organism needs to perform when making Bayesian inferences*” (Gigerenzer y Hoffrage, 1995). Si los algoritmos cognitivos de nuestros participantes fueron “diseñados” para manejar información en forma de frecuencias, es difícil explicar por qué el rendimiento no fue mayor al resolver nuestras tareas con frecuencias (ver cuadro 5).

Además, deberíamos recordar que nuestro resultado más relevante no es que el formato de representación no tuviera influencia en nuestros resultados, sino que hubo otra variable que sí la tuvo. El rendimiento era significativamente mejor cuando había correspondencia (de escala) entre las instrucciones y la respuesta que debían dar (comparado con cuando no la había). Por lo tanto, no se puede argumentar que nuestros problemas fueran demasiado simples computacionalmente como para que apareciera algún efecto. Había un efecto significativo pero el formato de representación y en particular las frecuencias no estaban implicadas.

Una segunda alegación en defensa de la aproximación frecuentista podría ser que nuestras frecuencias naturales no eran genuinas frecuencias naturales, o su argumento complementario, que nuestras probabilidades eran frecuencias naturales veladas. En principio, nuestra frase “de cada diez pimientos, 1 es picante y 9 no” cumple todos los criterios de las frecuencias naturales, aunque también podría ser argumentado que las frecuencias simples (aquellas que reflejan probabilidades simples) no son frecuencias naturales. Sin embargo, esto estaría contradiciendo el origen evolutivo atribuido a las frecuencias naturales. No hay muchas cosas sobre el pasado de nuestro sistema cognitivo que se puedan dar por sentado. No obstante, es una verdad analítica que, como hemos comentado antes, cada vez que uno de nuestros antepasados se encontró con las frecuencias necesarias para calcular una probabilidad Bayesiana, debió enfrentarse previamente a dos o tres probabilidades o frecuencias simples, ya que para el cálculo Bayesiano son necesarias varias probabilidades más simples (ver cuadro 5).

Cuadro 5. Cálculo de probabilidad Bayesiana.

**Cálculo de probabilidad Bayesiana**

En uno de los ejemplos clásicos más usados en la literatura (Gigerenzer et al., 1995) nos preguntan sobre p(Hipótesis|Dato), la probabilidad de que una mujer que ha dado positivo en una mamografía (Dato) tenga realmente cáncer de mama (Hipótesis).

Para realizar este cálculo necesitamos saber:

1. La probabilidad de que una mujer tenga cáncer de mama p(H).
2. La probabilidad de que una mujer con cáncer de mama dé positivo en una mamografía p(D|H).
3. La probabilidad de que una mujer sin cáncer de mama dé también positivo en una mamografía p(D|~H).

Con estos tres datos podemos calcular la probabilidad posterior Bayesiana de p(H|D) o mejor, p(Cáncer|Mamografía positiva):



(\*) p(D|H) se podría considerar como p(D) una vez situados en H. De hecho, la información en los problemas clásicos se aporta de esta manera (e.g. dentro del 1% con cáncer de mama H, un 80% dan positivo en la mamografía p(D)).

En el experimento 1 usamos un formato para expresar probabilidades que ha sido cuestionado ampliamente con el argumento de que la gente podría convertir éstos a frecuencias naturales con facilidad por la presencia de la categoría complementaria. Esto podría haber explicado la ausencia de diferencias, pero no explicaría el efecto encontrado a favor de las probabilidades y, como hemos visto en la variable *CC Explícita vs. Implícita* del experimento 3, la presencia de la categoría complementaria (pimientos no picantes en este caso), parece no tener una influencia determinante. En cualquier caso, pasamos el segundo experimento. En él, nuestras probabilidades son el epítome de la normalización no natural y, por lo tanto, no se puede argumentar que exista similitud alguna con las frecuencias naturales. A pesar de ello, de nuevo nuestros sujetos cometieron menos errores con probabilidades que con frecuencias. Por último, en el tercer experimento usamos una distribución de pimientos para la respuesta en lugar de pedirles a los participantes que predijeran un evento simple. En este caso, con una manipulación más compleja y teniendo en cuenta posibles factores que pudieran haber influido en los resultados anteriores, aparecieron diferencias, pero no en las direcciones esperadas. Se mostró como único factor significativo el *Emparejamiento I-R,* que no es más que la correspondencia (en escala) entre las instrucciones y la respuesta.

Mientras las frecuencias naturales sean presuntamente el formato de representación favorecido por la evolución, hay una explicación más bien complicada que el frente frecuentista debería dar. ¿Cómo puede un factor contextual como la correspondencia (en escala) entre instrucciones y respuesta tener un efecto más amplio que el del formato de representación? Si este simple hecho puede determinar en mayor medida los resultados, la conclusión es evidente. La complejidad computacional añadida al tener que realizar una sencilla operación de cambio de escala (pasar “1 de cada 10” a “10 de cada 100”) es mayor que la hipotética facilitación del algoritmo o módulo cognitivo especializado para el tratamiento de las frecuencias naturales, al menos con probabilidades simples.

Por último, el patrón de resultados de los pasados tres experimentos nos muestra la maleabilidad de los resultados en este tipo de problema al modificar varios factores contextuales.

Igual que pasaba en el primer experimento, los resultados globales de esta primera parte pueden ser explicados a partir de principios que nada tienen que ver con la superioridad de las frecuencias naturales. Por ejemplo, de acuerdo con la teoría de Modelos Mentales (Johnson-Laird et al., 1999), los problemas de frecuencias podrían ser representados como:

Pimiento picante (1)

Pimiento no picante (9)

Para las instrucciones de probabilidades, la representación podría ser:

Pimiento picante (10%)

Pimiento no picante (90%)

Cuando se les dice a los participantes que imaginen “un plato con diez de estos pimientos”, lo más probable, en el caso de la condición de frecuencias naturales, es que realicen una correspondencia directa entre la representación formada sobre la descripción previa, y la descripción desplato particular, porque los números son idénticos. Esto favorece la aparición de la *falacia del jugador* en la pregunta sobre el último pimiento del plato ya que el pimiento que queda se empareja con el pimiento picante de la descripción. Cuando la correspondencia es menos directa, como en la condición de probabilidades, los errores se reducen porque, al tener que manipular la representación, se evidencian las diferencias. Hay varias maneras en las que este emparejamiento se puede acentuar. Hemos comentado la primera antes, la correspondencia directa entre las instrucciones y su representación, aunque no se ha de confundir con la variable de *emparejamiento entre instrucciones y respuesta* que manipulábamos en el experimento 3*.* De hecho, ambas variables afectan de manera opuesta ya que, mientras la correspondencia entre los modelos mentales creados y las instrucciones agrava la incidencia de la *falacia del jugador*, aumentando los errores, el emparejamiento en escala entre instrucciones y respuesta reduce los pasos de cálculo necesarios facilitando la computación, además de evitar la correspondencia entre modelos e instrucciones. La segunda manera de acentuar el emparejamiento a nivel de modelos mentales es la presencia de la categoría complementaria, los pimientos no picantes. Esto es así porque, por ejemplo, en el caso que representamos más arriba, al preguntar por el último pimiento del plato cuando los otros nueve no fueron picantes, se establece una relación directa con los imaginados, quedando únicamente un solitario pimiento picante. Vemos esto especialmente en el caso de las frecuencias naturales, donde la diferencia entre situaciones donde hay más interferencia (*No Emparejamiento I-R* y *CC Explícita*) y menos (*Emparejamiento I-R y CC Implícita*) es especialmente grande (27%).

## 2. Cálculo de probabilidad Bayesiano, complejidad y formatos de representación

Hemos visto en el bloque anterior cómo con probabilidades simples la teórica superioridad de las frecuencias naturales se desvanecía, cuando no se mostraba como un efecto adverso, dejando como único factor relevante la complejidad añadida de un cambio de escala. Antes comentábamos que puede ser criticado el uso de probabilidades simples para falsar algo que en la literatura se ha trabajado casi en exclusiva con problemas complejos. En esta segunda parte vamos a intentar averiguar si las conclusiones a las que llegamos con probabilidades simples son extrapolables a los problemas clásicos. Queremos ver cuál es el papel real de la complejidad computacional en la resolución de los problemas Bayesianos en lo que se refiere a la aplicación del algoritmo Bayesiano.

Hay un problema fundamental que lastra las visiones frecuentistas del cálculo de probabilidades, el solapamiento de predicciones con otras alternativas basadas en diferencias en complejidad. Como ya hemos comentado, los frecuentistas (Brase y Barbey, 2006; Cosmides y Tooby, 1996; Gigerenzer y Hoffrage, 1995) proponen dos mecanismos a partir de los cuales se explican los mejores resultados de las frecuencias en problemas Bayesianos. El primer mecanismo es apenas controvertido, se refiere a cuestiones formales y es que el algoritmo Bayesiano es computacionalmente más sencillo cuando la información se presenta en frecuencias naturales. Esto es así simplemente porque su resolución implica menos pasos de cálculo, además de pasos menos complejos por el uso de números naturales en lugar de fracciones (Gigerenzer y Hoffrage, 1995). Aunque, como vimos antes, ellos son inconsistentes en el uso de este principio y, además, existe solapamiento entre los pasos de cálculo de los distintos menús de la información para frecuencias naturales.

El segundo mecanismo es evolutivo y ha generado un inmenso debate (Cosmides y Tooby, 1996; Gigerenzer, 2002; Gigerenzer y Hoffrage, 1995; Gigerenzer et al., 1991). Según ellos, la presencia de las frecuencias naturales a lo largo de la historia filogenética del ser humano ha contribuido a nuestra especialización y a la consiguiente existencia de un algoritmo cognitivo especialmente adaptado para éstas (Cosmides y Tooby, 1996; Gigerenzer y Fiedler, 2003; Gigerenzer y Hoffrage, 1995).

Es difícil hablar sobre el segundo de estos mecanismos, el evolutivo, por la dificultad que entraña la falsación empírica de este tipo de hipótesis (Andrews, Gangestad, y Matthews, 2003; Gould y Lewontin, 1979; Johnson-Laird et al., 1999; Lewontin, 1978, 1979), ya que la aproximación que se haga ha de ser por fuerza indirecta. Sí podemos, no obstante, hablar sobre el primer mecanismo, aunque, a priori, podría parecer que el acuerdo generalizado que existe sobre la mayor complejidad de los problemas Bayesianos en formato probabilístico hace que parezca casi irrelevante su tratamiento. Para aclarar este último extremo y con el objetivo último de determinar si existe algún tipo de conflicto entre ambos mecanismos postulados por los frecuentistas vamos a intentar entender mejor por qué nos dicen que las frecuencias naturales son tan sencillas en comparación con las probabilidades en problemas Bayesianos. Veamos por separado la tarea clásica (a partir de adaptación de tarea original de Eddy, 1982) usada por éstos para probabilidades y frecuencias naturales, así como una representación gráfica y la manera de resolver los problemas:

**Formato de Probabilidad Estándar**

Cuadro 6. Material experimental clásico usado por Gigerenzer y Hoffrage 1995 en la condición de probabilidades.

La probabilidad de cáncer de mama es de un 1% para mujeres de 40 años que participan en un chequeo rutinario.

Si una mujer tiene cáncer de mama, la probabilidad de que la mamografía dé positivo es de un 80%. Si una mujer no tiene cáncer de mama, la probabilidad de que la mamografía también dé positivo es de un 9,6%.

Una mujer de este grupo de edad da positivo en una mamografía en un chequeo rutinario. ¿Cuál es la probabilidad de que ella tenga realmente cáncer de mama? \_\_\_\_%

Esquema 1. Representación gráfica de la condición de probabilidades (en azul oscuro la información dada en el problema).

Cuadro 7. Resolución formal del problema en formato de probabilidades.

**Resolución de Formato de Probabilidad Estándar**

P (Cáncer) = 0,01

P (no Cáncer) = 0,99

P (Mamografía + | Cáncer) = 0,80

P (Mamografía + | no Cáncer) = 0,096



**Formato de Frecuencia Estándar**

Cuadro 8. Material experimental clásico usado por Gigerenzer y Hoffrage 1995 en la condición de frecuencias naturales.

Diez de cada 1000 mujeres de 40 años que participan en chequeos rutinarios tienen cáncer de mama.

Ocho de cada 10 mujeres con cáncer de mama darán positivo en una mamografía. De cada 990 mujeres sin cáncer de mama hay también 95 que darán positivo en una mamografía.

Aquí esta una nueva muestra representativa de mujeres de 40 años que han dado positivo en una mamografía en chequeos rutinarios. ¿Cuántas de estas mujeres esperas que tengan realmente cáncer de mama? \_\_\_\_ de \_\_\_\_

Esquema 2. Representación gráfica de la condición de frecuencias naturales (en azul oscuro la información dada en el problema).

Cuadro 9. Resolución formal del problema en formato de frecuencias naturales.

**Resolución de Formato de Frecuencia Estándar**

Cáncer ^ Mamografía + = 8 casos

no Cáncer ^ Mamografía + = 95 casos



Como se puede apreciar en el cuadro siguiente, en el caso de la condición de probabilidades se han de realizar un total de cuatro operaciones (cinco si se cuenta dos veces una de ellas) con números decimales para resolver el problema. En el caso de las frecuencias, la cantidad de operaciones se reduce a dos, y con números enteros.

Cuadro 10. Resolución formal de los problemas clásicos de Gigerenzer y Hoffrage 1995.

**Probabilidad**



1) (0,01) \* (0,8) = 0,008

2) ~~(0,01) \* (0,8) = 0,008~~

3) (0,99)\*(0,096) = 0,095

4) 0,008 + 0,095 = 0,103

**5) 0,008 / 0,103 = 0,077**

**Frecuencias Naturales**



1) (8) + (95) = 103

**2) (8) / 103 = 0,077**

Como decíamos, el problema surge, desde nuestro punto de vista, por el solapamiento entre la hipótesis evolucionista y la de la complejidad computacional. Es decir, a partir de los buenos resultados obtenidos con frecuencias naturales en problemas Bayesianos y basándose en la idea de que los dominios cognitivos para los que la mente ha sido diseñada a lo largo de la evolución pueden ser inferidos a partir de las regularidades en la estructura del mundo actual y el entorno (hipotético) en el que los humanos evolucionamos (Brase y Barbey, 2006), los frecuentistas (Cosmides y Tooby, 1996; Gigerenzer y Hoffrage, 1995) llegan a la conclusión de que los problemas representados con frecuencias naturales se resuelven con más facilidad porque la evolución las ha favorecido.

Creemos que la prueba de la superioridad de algún formato de representación sobre otro debería venir de datos empíricos en los que de forma no ambigua se separen las distintas variables que contribuyen hipotéticamente al resultado. Esa es la motivación de la siguiente serie de experimentos. En ella, por un lado, manipulamos de distintas maneras la complejidad de los problemas para intentar determinar cuál es el motivo final de las diferencias existentes y, por otro, controlamos factores cognitivos para tratar de averiguar hasta qué punto se puede hablar de una especialización cognitiva para la resolución de problemas de cálculo de probabilidad usando frecuencias naturales o, alternativamente, su resolución depende de habilidades cognitivas generales.

Finalmente, matizamos la postura frecuentista con respecto al papel de la complejidad computacional o aritmética, y es que, como hemos dicho antes, el uso que hacen ellos de ésta es inconsistente, ya que parecen no aplicarla a las diferencias entre los distintos menús de la información dentro del formato de frecuencias naturales.

En concreto, predecimos que al igualar las probabilidades previas por un lado y al usar clases de referencia absolutas por otro lado y, por lo tanto, en ambos casos, disminuir la dificultad en la condición de probabilidades, se reducirán las diferencias de rendimiento con las frecuencias naturales. También esperamos que la manipulación de los pasos de cálculo nos permita controlar el rendimiento de los sujetos de manera muy precisa, mucho más de lo que cuestiones como menú de la información y la versión descafeinada de complejidad usada por los frecuentistas permiten.

### 2.1. Manipulación de probabilidades previas

Queremos averiguar si los efectos que descubrimos con problemas basados en probabilidades simples son extrapolables al cálculo de probabilidad Bayesiano. Nos interesa comprobar, fundamentalmente, si igualando la complejidad computacional de los formatos frecuentista y probabilista desaparecerán las diferencias.

Pretendemos con este experimento, por un lado, replicar el experimento clásico de Gigerenzer y Hoffrage (1995) y, por otro, comparar el rendimiento de los participantes en función de las diferencias en formato de representación en situaciones con una complejidad computacional similar. El cambio esencial que introdujimos consistió en igualar las probabilidades previas para enfermos y no enfermos, reduciendo de este modo las diferencias en complejidad computacional en la resolución de los problemas.

En el cuadro 11 podemos ver la manera de resolver los problemas del tipo presentado en este experimento y cómo las diferencias en complejidad algorítmica de ambos problemas se reducen en comparación con los problemas clásicos que vimos más arriba.

Los frecuentistas han mostrado en innumerables ocasiones cómo el rendimiento de los participantes es mucho mejor con frecuencias naturales que con probabilidades en esta misma tarea. En el experimento que nos ocupa, el único cambio que realizamos en la nueva variante fue numérico, igualando las probabilidades previas de los distintos elementos del problema y redondeando, para evitar números decimales, lo que no debería modificar los resultados o, al menos, no hacer que desaparecieran las diferencias, en el caso en el que el ser humano estuviera mejor preparado para trabajar con frecuencias naturales.

Cuadro 11. Resolución formal de los problemas presentados en el experimento.

Para realizar este cálculo necesitamos saber:

1. La probabilidad a priori de que un hombre tenga la enfermedad p(H) y su complementaria p(~H).
2. La probabilidad de que un hombre con la enfermedad dé positivo en el test p(D|H)
3. La probabilidad de que un hombre sin la enfermedad dé también positivo en el test p(D|~H).

Con estos tres datos podemos calcular la probabilidad posterior Bayesiana **(1)** de p(H|D) o mejor, p(Enfermedad|Test positivo):

**(1)** 

Planteado en términos más sencillos **(2)**, lo que hacemos es dividir P(H) \* p(D|H), que es la probabilidad de que un hombre con la enfermedad dé positivo en el test, entre la probabilidad de p(D), que es la probabilidad de que un hombre dé positivo en el test (tenga o no tenga la enfermedad, p(H) \* p(D|H) + P(~H) \* p(D|~H)).

**(2)** 

En el enunciado del problema la información que se aporta es:

p(H) = 50% / 5000

p(D|H) = 100% / 5000

p(D|~H) = 1% / 50

Por lo que para resolverlo en las distintas condiciones el cálculo a realizar es:

**Probabilidades**

**(1)** 

**Frecuencias Naturales**

**(1)** 

La predicción que pretendemos falsar, por lo tanto, sería que los resultados serán superiores cuando los datos son representados con frecuencias naturales, aunque la complejidad algorítmica sea similar. Si por el contrario, en condiciones de igualdad computacional o algorítmica no hubiera diferencias, se necesitarían más datos para sostener la hipótesis evolutiva. En este cuarto experimento estaríamos poniendo a prueba de manera algo simplista la influencia del algoritmo o módulo especializado en frecuencias en comparación con el peso de la complejidad algorítmica o aritmética.

Los sujetos debían calcular la probabilidad de que un hombre que diera positivo en un test experimental tuviera realmente la enfermedad para la cual se estaba probando el test. Antes de los datos experimentales había unas instrucciones genéricas sobre la tarea con la finalidad de hacer más sencilla su comprensión y eliminar cualquier dificultad añadida a la hora de computar la respuesta.

#### Método

#### Participantes y diseño

Un total de 100 estudiantes de psicología de la Universidad de Murcia completaron un cuestionario en clase. Fueron asignados de forma aleatoria a alguno de los cuatro grupos correspondientes al factor entre-sujetos, *Formato de representación* (frecuencias/probabilidades) en cada una de las dos variantes (Problema clásico/ Probabilidades previas igualadas).

Tabla 9. Variables usadas y representación esquemática del material.

**Variable independiente**

Entre-sujetos:

* Formato de representación
  + Probabilidades (0,01%)
  + Frecuencias (1 de cada 10000)

Variantes

* Problema clásico
* Probabilidades previas igualadas

Nuestro objetivo era buscar diferencias en *Formato de representación* en las variantes de *Problema clásico* y *Probabilidades previas igualadas* por separado. La variante clásica del problema fue usada como control de que los sujetos realizaban correctamente la tarea y de que podía reproducirse el efecto clásico con nuestra muestra y, por lo tanto, cualquier ausencia de diferencias no sería por falta de capacidad discriminativa en el experimento.

#### Procedimiento y materiales

Los participantes contestaron a un cuestionario en clase que incluía un problema con una pregunta relativa a la probabilidad de que un hombre que diera positivo en un test experimental tuviera realmente una enfermedad.

En la tabla siguiente se puede ver el material experimental:

Tabla 10. Material del experimento 4 (en cursiva)

***Instrucciones***

*Se está probando un test para intentar detectar la presencia de una enfermedad. Presentamos a continuación datos relativos a la gente infectada por la enfermedad y las tasas de error y acierto del test.*

*Tienes que leer muy atentamente el enunciado que se te presenta, entenderlo y contestar a la pregunta.*

*Puedes hacer cálculos, esquemas, dibujos y todo lo que necesites para ayudarte a entender y resolver el problema. De hecho, nos interesa especialmente esa información.*

*Toma todo el tiempo que necesites.*

* Probabilidades
  + Problema clásico [Respuesta correcta = 50%]

***Problema***

*Un 0,01% de los hombres están infectados por el virus y darán positivo en el test.*

*Del 99,99% que no están infectados, un 0,01% dará también positivo en el test.*

*¿Cuál es la probabilidad de que un hombre que da positivo en el test tenga realmente el virus? \_\_\_\_\_%*

* + Probabilidades previas igualadas [Respuesta correcta = 99%]

***Problema***

*Un 50% de los hombres están infectados por el virus y darán positivo en el test.*

*Del 50% que no están infectados, un 1% dará también positivo en el test.*

*¿Cuál es la probabilidad de que un hombre que da positivo en el test tenga realmente el virus? \_\_\_\_\_%*

* Frecuencias
  + Problema clásico [Respuesta correcta = 50%]

***Problema***

*Imagina a 10000 hombres.*

*1 está infectado por el virus y dará positivo en el test.*

*De los 9999 que no están infectados, 1 dará también positivo en el test.*

*¿Cuantos de los hombres con resultados positivos en el test esperas que realmente tengan el virus? \_\_\_\_\_ de cada \_\_\_\_\_*

* + Probabilidades previas igualadas [Respuesta correcta = 99%]

***Problema***

*Imagina a 10000 hombres.*

*5000 están infectados por el virus y darán positivo en el test.*

*De los 5000 que no están infectados, 50 darán también positivo en el test.*

*¿Cuantos de los hombres con resultados positivos en el test esperas que realmente tengan el virus? \_\_\_\_\_ de cada \_\_\_\_\_*

El tercer párrafo de las instrucciones, el relativo a cálculos, esquemas y demás, se incluyó con la finalidad de dar mayor libertad y potenciar un procesamiento más profundo del problema. No analizamos de manera cualitativa esta información por la gran cantidad de sujetos que no realizaron anotaciones.

#### Resultados

Los resultados se presentan en la tabla 11. Nuestro objetivo en este caso era comparar por separado las versiones *Problema clásico* y *Probabilidades previas igualadas*. Como comentábamos antes, la versión clásica la usamos como control con la intención de comprobar que los sujetos estaban resolviendo la tarea correctamente.

En la versión clásica el ANOVA mostró un efecto significativo del *Formato de representación* (F(1, 49)= 13,486, MSE= 0,148, p= 0,001, η2=0,22), respondiendo mejor los participantes en la condición de Frecuencias (44% aciertos) que en la de Probabilidades (4% aciertos). Este resultado corresponde al efecto clásico encontrado por Gigerenzer y Hoffrage (1995), entre otros, y ratifica que nuestros sujetos están resolviendo la tarea de igual manera que en los estudios clásicos.

En el caso de la versión con probabilidades previas igualadas, el ANOVA no mostró diferencias significativas en la variable *Formato de representación* (F(1, 49)= 0,915, MSE= 0,197, p= 0,344, Potencia=0,155). Los promedios de aciertos para Frecuencias (32% aciertos, Intervalo de confianza= 14,2/49,8) y Probabilidades (20% aciertos, Intervalo de confianza= 2,2/37,8) no son, por lo tanto, significativamente diferentes y, de hecho, la varianza explicada por el error es mayor que la explicada por el efecto (F<1).

Tabla 11. Porcentaje de respuestas correctas y (entre paréntesis) el error promedio (probabilidad real – probabilidad calculada) de las respuestas de los participantes para cada condición experimental.

**Aciertos Intervalo de Confianza (95%) N**

**Problema clásico**

**Frecuencias** 44 (28) 28,5 / 59,5 25

**Probabilidades** 4 (40) -11,5 / 19,5 25

**Probabilidades previas igualadas**

**Frecuencias** 32 (33) 14,2 / 49,8 25

**Probabilidades** 20 (34) 2,2 / 37,8 25

En este cuarto experimento tenemos, por un lado, una réplica del experimento clásico que muestra una superioridad más que evidente de las Frecuencias naturales sobre las Probabilidades y, por otro lado, una versión modificada del mismo en la que, al igualar las probabilidades previas de cada uno de los grupos (Enfermo/no Enfermo), estas diferencias se reducen hasta el punto de perder valor estadístico.

A pesar de los problemas que conlleva cualquier afirmación realizada a partir de la ausencia de diferencias, nos gustaría destacar que la hipótesis que guió la manipulación que realizamos es congruente con esta ausencia de diferencias y, además, la potencia es suficiente para conseguir un efecto del tamaño existente en la versión clásica del problema. Así pues, estos resultados van en la dirección esperada, apoyando la hipótesis de la complejidad computacional como el único factor relevante a la hora de explicar las diferencias clásicas obtenidas en cálculo de probabilidad con distintos formatos de representación.

### 2.2. Clases de referencia y formatos de representación

Por lo que hemos visto en los experimentos anteriores, parece que la complejidad aritmética o algorítmica de los problemas, y cuestiones contextuales como la coincidencia entre las instrucciones y el formato de respuesta, son factores que habría que tener en cuenta a la hora de resolver el puzle de los formatos de representación en el cálculo de probabilidad. Si se analiza con más detalle el tipo de tareas de cálculo Bayesiano empleadas por Gigerenzer y Hoffrage (1995), entre otros, se puede ver que hay algunas dimensiones importantes que requieren una mayor atención. Autores como Macchi hablan de la estructura partitiva (una formulación que según sus proponentes clarifica las relaciones entre los distintos subconjuntos identificando el grupo de referencia, reduciendo la confusión, mostrando la independencia de los datos y permitiendo detectar y hacer posible el uso de relaciones entre los datos) como elemento explicativo (Macchi, 2000; Macchi y Mosconi, 1998), aunque son duramente criticados por Gigerenzer y otros (Hoffrage et al., 2002) en base a dos cuestiones fundamentales. En primer lugar, según Hoffrage y colaboradores, Macchi muestra que las frecuencias normalizadas no facilitan el razonamiento, a partir de lo que concluye que el efecto facilitatorio no es debido a las frecuencias, lo que es especialmente grave, ya que los frecuentistas explícitamente indican que la normalización es un rasgo característico de las probabilidades (Gigerenzer y Hoffrage, 1995). En segundo lugar, dicen que la presencia de la muestra total (*total sample*) supone un acercamiento al formato frecuentista y, por lo tanto, una injusta ventaja para las probabilidades. También Girotto y Gonzalez (Girotto y Gonzalez, 2001, 2002) pugnan con Gigerenzer, Hoffrage y colaboradores (Hoffrage et al., 2002), sugiriendo que los individuos son perfectamente capaces de razonar probabilísticamente, siempre y cuando puedan apoyarse en una representación que muestre claramente los subconjuntos de probabilidades. De nuevo, el frente frecuentista contesta alegando que lo único que Girotto y Gonzalez están haciendo es redescubrir las propiedades de las frecuencias naturales ya indicadas por ellos, dándoles otro nombre.

Finalmente, Fiedler y colaboradores (Fiedler et al., 2000) aluden a la clase o set de referencia a la que se refiere la información, estableciendo una elegante diferencia entre las situaciones en las que ésta es común para todos los datos y aquellas en las que la clase de referencia de la información se refiere a subconjuntos parciales. Cómo no, son también criticados por Hoffrage y colaboradores (Hoffrage et al., 2002) en base a la presencia de la muestra total como posible elemento facilitador para las condiciones probabilistas, obviando las preguntas planteadas por las llamativas contradicciones con la hipótesis frecuentista y sin aportar datos empíricos sobre la influencia real de la presencia o ausencia de la muestra total.

Todos estos factores, principalmente extensionales, han generado un extraordinario debate. Hay sobre algunos de ellos acuerdos parciales pero, sobre todo, siguen existiendo problemas de definición y cuestiones no comprobadas empíricamente.

En este experimento realizamos una réplica, y al tiempo, versión, del experimento de Fiedler (2000), centrándonos en los factores que consideramos más importantes (*Formato de representación* y *Complejidad*) y añadiendo la variable *Muestra total* con la intención de finalmente tener pruebas empíricas de su influencia y, al mismo tiempo, intentar contestar a la crítica hecha por los frecuentistas de la que hablamos más arriba y que dice que la presencia de la muestra total hace sencilla la conversión a frecuencias naturales de las probabilidades.

Además de lo anterior, queremos evidenciar que la comparación clásica de la que se extrae la superioridad de las frecuencias naturales, y que se realiza entre probabilidades relativas y frecuencias absolutas, es injusta. Veremos más adelante por qué y qué significa esto exactamente, pero el simple hecho de no controlar la clase de representación de la información (absoluta/relativa), en el caso de que esta variable sea relevante, minimiza la importancia de los resultados clásicos.

#### Método

#### Participantes y diseño

Un total de 351 estudiantes de psicología de la Katholieke Universiteit Leuven, en Bélgica, completaron un cuestionario en flamenco como parte de una serie de procedimientos experimentales necesarios para aprobar una asignatura obligatoria. Fueron asignados de forma aleatoria a alguno de los ocho grupos resultantes de los factores entre-sujetos, *Formato de representación* (Frecuencias/Probabilidades), *Complejidad* (Absoluta/Relativa) y *Muestra total* (Presencia de muestra tota/Ausencia de muestra total).

Tabla 12. Variables usadas y representación esquemática del material.

**Variables independientes**

Entre-sujetos:

* Formato de representación
  + Probabilidades
  + Frecuencias
* Complejidad
  + Absoluta
  + Relativa
* Muestra total
  + Presencia de muestra total
  + Ausencia de muestra total

#### Procedimiento y materiales

Los participantes contestaron a una sola pregunta relativa a la probabilidad de que una mujer tuviera cáncer de mama dado que el resultado de la mamografía había sido positivo. Pasaron, además, como parte de un amplio protocolo experimental, una prueba de carácter cognitivo, el Cognitive Reflection Test (CRT).

En la tabla siguiente se puede ver el material experimental y en la tabla 14, una representación esquemática del material experimental para su mejor comprensión:

Tabla 13. Material del experimento 5 (en cursiva).

*Intenta resolver el siguiente problema de razonamiento sobre mamografías. La mamografía es una prueba médica que nos permite detectar pequeños bultos en las mamas (un resultado positivo indica la presencia de bultos que pueden ser o no cancerígenos). Este método se usa para ayudar a diagnosticar de forma temprana el cáncer de mama.*

*Lee la información del problema atentamente y escribe tu respuesta. Puedes tomar todo el tiempo que quieras para resolver el problema.*

* Probabilidades
  + Absoluta
    - Presencia de muestra total

*Se realizó un estudio a partir de datos de 1000 mujeres.*

*El 89,5% de las mujeres no tenía cáncer de mama y la mamografía le dio negativo.*

*El 9,5% de las mujeres no tenía cáncer de mama y la mamografía le dio positivo.*

*El 0,8% de las mujeres tenía cáncer de mama y la mamografía le dio positivo, y el 0,2% de las mujeres tenía cáncer de mama y la mamografía le dio negativo.*

* + - Ausencia de muestra total

*Se realizó un estudio a partir de datos de un gran número de mujeres.*

*El 89,5% de las mujeres no tenía cáncer de mama y la mamografía le dio negativo.*

*El 9,5% de las mujeres no tenía cáncer de mama y la mamografía le dio positivo.*

*El 0,8% de las mujeres tenía cáncer de mama y la mamografía le dio positivo, y el 0,2% de las mujeres tenía cáncer de mama y la mamografía le dio negativo.*

* + Relativa
    - Presencia de muestra total

*Se realizó un estudio a partir de datos de 1000 mujeres.*

*El 99% de las mujeres no tenía cáncer de mama y el 1% tenía cáncer de mama.*

*De las mujeres sin cáncer de mama, a un 10% la mamografía le dio positivo y a un 90% negativo.*

*Y de las mujeres con cáncer de mama, a un 80% la mamografía le dio positivo y a un 20% negativo.*

* + - Ausencia de muestra total

*Se realizó un estudio a partir de datos de un gran número de mujeres.*

*El 99% de las mujeres no tenía cáncer de mama y el 1% tenía cáncer de mama.*

*De las mujeres sin cáncer de mama, a un 10% la mamografía le dio positivo y a un 90% negativo.*

*Y de las mujeres con cáncer de mama, a un 80% la mamografía le dio positivo y a un 20% negativo.*

* Frecuencias
  + Absoluta
    - Presencia de muestra total

*Se realizó un estudio a partir de datos de 1000 mujeres.*

*895 mujeres no tenían cáncer de mama y la mamografía les dio negativo.*

*95 mujeres no tenían cáncer de mama y la mamografía les dio positivo.*

*8 mujeres tenían cáncer de mama y la mamografía les dio positivo, y 2 mujeres tenían cáncer de mama y la mamografía les dio negativo.*

* + - Ausencia de muestra total

*Se realizó un estudio a partir de datos de un gran número de mujeres.*

*895 mujeres no tenían cáncer de mama y la mamografía les dio negativo.*

*95 mujeres no tenían cáncer de mama y la mamografía les dio positivo.*

*8 mujeres tenían cáncer de mama y la mamografía les dio positivo, y 2 mujeres tenían cáncer de mama y la mamografía les dio negativo.*

* + Relativa
    - Presencia de muestra total

*Se realizó un estudio a partir de datos de 1000 mujeres.*

*990 mujeres no tenían cáncer de mama y 10 tenían cáncer de mama.*

*De las mujeres sin cáncer de mama, a 95 la mamografía le dio positivo y a 895 negativo.*

*Y de las mujeres con cáncer de mama, a 8 la mamografía le dio positivo y a 2 negativo.*

* + - Ausencia de muestra total

*Se realizó un estudio a partir de datos de un gran número de mujeres.*

*990 mujeres no tenían cáncer de mama y 10 tenían cáncer de mama.*

*De las mujeres sin cáncer de mama, a 95 la mamografía le dio positivo y a 895 negativo.*

*Y de las mujeres con cáncer de mama, a 8 la mamografía le dio positivo y a 2 negativo.*

***¿Cuál es la probabilidad de tener cáncer de mama si a una mujer la mamografía le ha dado positivo?***

*La probabilidad es\_\_\_\_\_\_\_\_\_\_\_\_\_* (da una respuesta numérica: por ejemplo, “La probabilidad es x%” o “La probabilidad es x de y”) *.*

Tabla 14. Representación esquemática del material del experimento obviando la del factor *Muestra total (en azul claro, la información no aportada)*.

**Probabilidades – Absoluta**

**Probabilidades – Relativa**

**Frecuencias – Absoluta**

**Frecuencias – Relativa**

La distinción de Fiedler referente a sets o clases de referencia es de gran utilidad para entender las dificultades con las que se encuentra la gente a la hora de resolver este tipo de problemas, pero es víctima de una inevitable dificultad. En el caso del formato de representación de probabilidades, la distinción funciona a la perfección. Al usar una posibilidad previa absoluta p(D&H), la dificultad computacional del problema se reduce de forma automática equiparándose con la de los problemas frecuentistas clásicos, ya que ésta incluye la probabilidad previa p(H), lo que simplifica la computación. Esto no es así con probabilidades relativas p(D|H), donde es necesario usar también la probabilidad previa p(H) para realizar el cálculo. Sin embargo, no pasa lo mismo con las frecuencias naturales con una clase de referencia relativa. Como podemos ver más arriba, en esta condición se aporta más información, pero la información crítica está ahí. Computacionalmente, la dificultad sigue siendo idéntica a la condición de frecuencias absolutas, no es necesario ningún cálculo extra, aunque sí es necesario un mayor esfuerzo de comprensión.

El Cognitive Reflection Test (CRT) (ver cuadro 12) es una sencilla prueba que refleja “*cognitive reflection*”, la habilidad o disposición a resistirse a informar de la primera respuesta que a uno le viene a la mente (Frederick, 2005). Los tres ítems que componen el CRT son sencillos en cuanto a que se entienden con facilidad al ser explicados, pero para llegar a la respuesta correcta, se ha de inhibir previamente la respuesta que, siendo errónea, parece intuitiva y automáticamente cierta.

Podríamos decir que el CRT mide de manera aproximada habilidades cognitivas generales, así como una cierta predisposición a la impulsividad. Se ha establecido una relación considerable entre puntuaciones en el CRT y medidas ampliamente usadas de capacidades cognitivas generales como el Scholastic Achievement Test (SAT) y el American College Test (ACT) (correlaciones de 0,44 y 0,46 respectivamente), lo que valida este test como medida de habilidades cognitivas generales.

Cuadro 12. Cognitive Reflection Test.

(1) Un bate y una pelota cuestan 1,10€ en total. El bate cuesta 1€ más que la pelota. ¿Cuánto cuesta la pelota? \_\_\_\_\_\_ céntimos.

(2) Si 5 máquinas tardan 5 minutos en hacer 5 aparatos, ¿cuánto tardarán 100 máquinas en hacer 100 aparatos? \_\_\_\_\_ minutos.

(3) En un lago, hay un grupo de nenúfares. Cada día, el grupo de nenúfares dobla su tamaño. Si son necesarios 48 días para que los nenúfares cubran todo el lago, ¿cuánto tardarán en cubrir la mitad del lago? \_\_\_\_\_ días.

#### Resultados

Los resultados se presentan en la tabla 15. El ANOVA mostró una interacción entre *Formato de representación* y *Complejidad* (F(1, 343)=9,233, MSE= 0,185, p= 0,003, η2=0,023), producida por los malos resultados en probabilidades relativas (1% aciertos) comparado con las demás condiciones.

Hay diferencias significativas en los efectos simples al comparar probabilidades relativas (1% aciertos) con frecuencias absolutas (46% aciertos), (t’(96,401)= 8,256 p< 0,0001), en lo que sería una replicación del efecto clásico encontrado por Gigerenzer y Hoffrage (1995) y lo visto en el experimento anterior. También se aprecia una diferencia significativa al compararlos con frecuencias relativas (37% aciertos), (t’(93,342)= 6,712 p<0,0001), y con probabilidades absolutas (38% aciertos), (t’(98,076)= 6,949 p< 0,0001).

Los resultados tan bajos en probabilidades relativas crean diferencias significativas en efectos principales, ya que arrastran hacia abajo la condición de probabilidades (20% aciertos) y, al compararla con frecuencias (42% aciertos), tenemos un efecto principal significativo de *Formato de representación* (F(1, 343)=23,419, MSE= 0,185, p< 0,0001, η2=0,058). Aparece también un efecto principal significativo de *Complejidad* (F(1, 343)=24,300, MSE= 0,185, p< 0,0001, η2=0,06), en el que el rendimiento de los sujetos es mucho peor en los problemas con una clase de referencia relativa (19% aciertos) en comparación con los problemas con clase de referencia absoluta (42% aciertos).

No hay, además, diferencias al comparar probabilidades absolutas (38% aciertos, Intervalo de confianza= 26,3/47,2) y frecuencias absolutas (46% aciertos, Intervalo de confianza= 36,5/57,6; F(1, 180)=1.864, MSE= 0,244, p= 0,174, Potencia=0,274), lo que apunta en la dirección de que la complejidad es más importante que el formato de representación. Finalmente, no fue significativo el efecto de *Muestra total*, donde la condición de Presencia de muestra total (30% aciertos) no difería de la de Ausencia de muestra total (32% aciertos)(F(1, 343)=0,184, MSE= 0,185, p= 0,669, Potencia=0,071), algo que, desde luego, no apoya su uso como elemento crítico que pueda contribuir a la similitud de las probabilidades con las frecuencias naturales, como han comentado Hoffrage y Gigerenzer (2002).

Tal y como explicábamos antes, si se analiza con detalle la condición de frecuencias relativas es fácil cuestionar su adecuación como condición relativa. Se puede manipular la redacción de la condición para darle un “aura” de relatividad, pero el número aportado será siempre el mismo: 8 de cada 1000 mujeres y 8 de cada 10 mujeres con cáncer se diferencian en la clase de referencia, pero en ambos casos el dato crítico es idéntico, 8. La complejidad computacional de esta condición es idéntica a la de frecuencias absolutas. Esto nos ayudó a interpretar la ausencia de diferencias entre frecuencias relativas (37% aciertos) y absolutas (46% aciertos), (t’(173)= -1,186, p= 0,237).

Tabla 15. Porcentajes de respuestas correctas y (entre paréntesis) la desviación estándar de cada condición experimental.

**Formato Complejidad Muestra total Aciertos N**

Frecuencias Relativas Ausencia de muestra total 34 (48) 44

Presencia de muestra total 40 (50) 42

Total 37 (49) 86

Absolutas Ausencia de muestra total 56 (50) 45

Presencia de muestra total 36 (49) 44

Total 46 (50) 89

Total Ausencia de muestra total 45 (50) 89

Presencia de muestra total 38 (49) 86

Total 42 (49) 175

Probabilidades Relativas Ausencia de muestra total 2 (15) 44

Presencia de muestra total 0 (0) 42

Total 1 (11) 86

Absolutas Ausencia de muestra total 34 (48) 44

Presencia de muestra total 41 (50) 46

Total 38 (49) 90

Total Ausencia de muestra total 18 (39) 88

Presencia de muestra total 22 (41) 88

Total 20 (40) 176

Total Relativas Ausencia de muestra total 18 (39) 88

Presencia de muestra total 20 (40) 84

Total 19 (39) 172

Absolutas Ausencia de muestra total 45 (50) 89

Presencia de muestra total 39 (49) 90

Total 42 (49) 179

Total Ausencia de muestra total 32 (47) 177

Presencia de muestra total 30 (46) 174

Total 31 (46) 351

Cuadro 13. Complejidad (en pasos de cálculo necesarios) de las distintas condiciones del experimento 5.

**Probabilidades**

**Absolutas**

Pasos de cálculo necesarios: 2



**Relativas**

Pasos de cálculo necesarios: 4/5



**Frecuencias Naturales**

**Absolutas**

Pasos de cálculo necesarios: 2



**Relativas**

Pasos de cálculo necesarios: 2



Vemos, por lo tanto, que el formato de representación pierde importancia cuando se tiene en cuenta la clase de referencia de la información, que este último factor resulta crítico a la hora de evaluar el rendimiento de los sujetos con problemas Bayesianos y que, finalmente, la presencia o ausencia de la muestra total no parece determinante.

#### Algoritmos cognitivos y CRT

Como dijimos más arriba, es realmente difícil comprobar empíricamente las afirmaciones relativas a aspectos evolutivos, especialmente a aquellos referidos a habilidades cognitivas (Andrews et al., 2003; Gould y Lewontin, 1979; Lewontin, 1978, 1979), y este caso no es una excepción. Las aproximaciones han de ser por fuerza de carácter indirecto y su valor puede ser cuestionado con más facilidad.

En el desarrollo de este trabajo nos basamos en un principio muy simple, contraponemos una tradición que ve la mente como una máquina de procesamiento de información de carácter general con otra que describe la mente como un conjunto de habilidades, cada una de las cuales ha sido diseñada por selección natural para lidiar con aspectos específicos del mundo (Brase y Barbey, 2006; Gigerenzer y Fiedler, 2003; Todd y Gigerenzer, 2000, 2003, 2007). Esta última visión da a entender, desde nuestro punto de vista, que el rendimiento en problemas relacionados con el procesamiento de este algoritmo o módulo especializado debería ser independiente de otros algoritmos cognitivos y, especialmente, de inteligencia general. Por lo tanto, en el caso de resolverse este tipo de problemas a partir de un sistema general de cómputo, cuantos más recursos haya disponibles más posible es que se dé la respuesta correcta. Por el contrario, si se usa un módulo automático y encapsulado, los recursos cognitivos generales no deberían tener especial influencia sobre los resultados (De Neys, 2007).

Dicho de manera más clara, si a lo largo de nuestra historia evolutiva ha habido presiones selectivas que favorecieron la aparición de un algoritmo cognitivo o módulo especializado y aislado para el tratamiento de frecuencias naturales, su propia especialización debería hacerlo, hasta cierto punto, independiente de inteligencia general.

Para comprobar este extremo dividimos nuestra muestra en dos grupos de acuerdo con las puntuaciones en el CRT (usando la mediana como criterio, 50% superior y 50% inferior). El rendimiento en la tarea clásica de frecuencias naturales fue muy superior en el grupo de superior del CRT comparado con el grupo inferior (69% aciertos en CRT superior y 26% en el grupo inferior) (F(1, 83)=19,4, MSE= 0,207, p< 0,0001, η2=0,19). En la tabla 16 se pueden ver los resultados para el resto de condiciones en función del rendimiento en el CRT. Aquí no analizamos las otras condiciones por su ausencia de interés para hablar sobre la hipótesis evolutiva.

Nuestros resultados muestran, además, una correlación general entre el CRT y los resultados en los problemas Bayesianos de 0,322 (p<0,001), de 0,252 (p=0,001) para los problemas Bayesianos en formato probabilístico y de 0,41 (p<0,001) para los problemas en formato frecuentista.

Tabla 16. Porcentajes de respuestas correctas y (entre paréntesis) la desviación estándar de cada condición experimental en los grupos de puntuaciones Altas y Bajas en el CRT.

**CRT Formato Complejidad Aciertos N**

50% inferior Frecuencias Relativas 20 (41) 44

Absolutas 26 (44) 43

Probabilidades Relativas 0 (0) 44

Absolutas 22 (42) 37

50% superior Frecuencias Relativas 55 (50) 40

Absolutas 69 (47) 42

Probabilidades Relativas 2 (16) 41

Absolutas 48 (51) 50

La fuerte relación entre una medida de rendimiento cognitivo general y los aciertos en esta tarea es difícil de explicar desde la perspectiva de la *Ecological Rationality*, ya que pone en duda la existencia de un módulo o algoritmo cognitivo específicamente preparado para trabajar con frecuencias naturales. Por otro lado, este tipo de resultado es congruente con explicaciones cercanas a visiones de la mente como una herramienta de propósito general.

### 2.3. Complejidad aritmética y formatos de representación

En el experimento anterior mostramos cómo la clase de referencia de la información y la diferencial complejidad aritmética asociada determina, en mayor medida que el formato de representación, los resultados en tareas de cálculo de probabilidad Bayesiana. Una parte de esta afirmación es congruente con la realizada por el frente frecuentista (Gigerenzer y Hoffrage, 1995), aunque necesitamos profundizar más para determinar con precisión qué es lo que subyace a las diferencias en rendimiento entre probabilidades y frecuencias naturales o, más bien, entre situaciones con clase de referencia absoluta y relativa.

Los frecuentistas aceptan que la menor complejidad computacional de las frecuencias naturales es responsable del mejor rendimiento de los sujetos con éstas. Crean, de manera más o menos explícita, una asociación exclusiva entre simplicidad computacional y frecuencias naturales. Ya hemos visto en el experimento anterior que no debería ser así puesto que la clase de referencia de la información es crítica para su correcta computación. Una manera de clarificar las causas subyacentes es detallar y manipular el número exacto de pasos de cálculo necesarios para resolver los distintos problemas y comprobar si este factor puede ser un predictor que nos permita explicar confusiones y datos contrapuestos existentes en la literatura.

Pretendemos, por lo tanto, comprobar si la complejidad aritmética, definida como el número de operaciones mínimas necesarias para resolver el problema, es una buena medida de dificultad y, por extensión, sirve para predecir el rendimiento a la hora de resolver problemas de cálculo de probabilidad Bayesiana.

Cuadro 14. Complejidad (en pasos de cálculo necesarios) de las distintas condiciones del experimento.

**Probabilidades**

**Pasos de cálculo necesarios: 4/5**



**Pasos de cálculo necesarios: 2**



**Pasos de cálculo necesarios: 0/1**



**Frecuencias Naturales**

**Pasos de cálculo necesarios: 2**



#### Método

#### Participantes y diseño

Un total de 124 estudiantes de bachillerato, en el día de las Pruebas de Acceso a la Universidad (PAU), completaron un cuestionario. Fueron asignados de forma aleatoria a alguno de los cuatro grupos correspondientes a los factores entre-sujetos, *Formato de representación* (frecuencias/probabilidades) y *Complejidad aritmética (pasos de cálculo necesarios)* (1, 2, 4).

Tabla 17. Variables usadas y representación esquemática del material.

**Variables independientes**

Entre-sujetos:

* Formato de representación
  + Probabilidades
    - Complejidad aritmética
  + 1 paso de cálculo
  + 2 pasos de cálculo
  + 4 pasos de cálculo
  + Frecuencias
    - Complejidad aritmética
  + 2 pasos de cálculo

#### Procedimiento y materiales

Los participantes contestaron a una sola pregunta relativa a la probabilidad de que una mujer tuviera cáncer de mama dado que el resultado de la mamografía había dado positivo.

En la tabla siguiente se puede ver el material experimental:

Tabla 18. Material del experimento (en cursiva).

***Instrucciones***

*Intenta resolver el siguiente problema de razonamiento sobre mamografías. La mamografía es una prueba médica que nos permite detectar pequeños bultos en las mamas (un resultado positivo indica la presencia de bultos que pueden ser o no cancerígenos). Este método se usa para ayudar a diagnosticar de forma temprana el cáncer de mama.*

*Lee la información del problema atentamente y escribe tu respuesta. Puedes tomar todo el tiempo que quieras para resolver el problema.*

*Se realiza un estudio a partir de datos de un gran número de mujeres.*

* Probabilidades
  + 1 paso de cálculo

*El 90% de las mujeres no tenía cáncer de mama y el 10% tenía cáncer de mama.*

*De las mujeres sin cáncer de mama, a un 0% la mamografía le dio positivo y a un 100%, negativo.*

*Y de las mujeres con cáncer de mama, a un 100% la mamografía le dio positivo y a un 0%, negativo.*

* + 2 pasos de cálculo

*El 90% de las mujeres no tenía cáncer de mama y el 10% tenía cáncer de mama.*

*De las mujeres sin cáncer de mama, a un 100% la mamografía le dio positivo y a un 0%, negativo.*

*Y de las mujeres con cáncer de mama, a un 100% la mamografía le dio positivo y a un 0%, negativo.*

* + 4 pasos de cálculo

*El 90% de las mujeres no tenía cáncer de mama y el 10% tenía cáncer de mama.*

*De las mujeres sin cáncer de mama, a un 10% la mamografía le dio positivo y a un 90%, negativo.*

*Y de las mujeres con cáncer de mama, a un 80% la mamografía le dio positivo y a un 20%, negativo.*

* Frecuencias
  + 2 pasos de cálculo

*810 mujeres no tenían cáncer de mama y la mamografía les dio negativo.*

*90 mujeres no tenían cáncer de mama y la mamografía les dio positivo.*

*80 mujeres tenían cáncer de mama y la mamografía les daba positivo y 20 mujeres tenían cáncer de mama y la mamografía les dio negativo.*

**¿Cuál es la probabilidad de tener cáncer de mama si a una mujer la mamografía le ha dado positivo?**

La probabilidad es\_\_\_\_\_\_\_\_\_\_\_\_\_.

A continuación, en la tabla 19, se puede ver una representación esquemática de las distintas condiciones experimentales, así como de la manera de resolverlas.Tabla 19. Representación esquemática del material del experimento 6 y pasos de cálculo necesarios para su resolución.

**Probabilidades – Complejidad aritmética = 4**

Pasos de cálculo: 4/5



**Probabilidades – Complejidad aritmética = 2**

Pasos de cálculo: 2



**Probabilidades – Complejidad aritmética = 1**

Pasos de cálculo: 1/0



**Frecuencias – Complejidad aritmética = 2**

Pasos de cálculo: 2



#### Resultados

Los resultados se presentan en la tabla 20. Como hemos dicho antes, pretendíamos comprobar si dentro de la condición de probabilidades la complejidad aritmética, definida como el número de operaciones mínimas necesarias para resolver el problema, es una buena medida de dificultad. Una vez hecho esto, nuestra intención era realizar la comparación entre la condición de probabilidades en sus diferentes niveles de complejidad con la tarea clásica usada por Gigerenzer y otros para tratar de definir con más claridad los motivos subyacentes a las diferencias en rendimiento.

Se definió a priori el orden de los niveles de dificultad a partir del número de pasos de cálculo necesarios pasa su resolución, quedando la complejidad de las diferentes condiciones de la siguiente forma: Complejidad aritmética 4 sería más difícil que Complejidad aritmética 2, que a su vez sería más difícil que Complejidad aritmética 1. Se utilizó este diseño en el ANOVA para verificar dicho modelo.

El ANOVA mostró un efecto significativo de la *Complejidad aritmética* (F(3, 121)= 24,684, MSE= 0,150, p< 0,0001, η2=0,29), ya que los participantes respondieron mejor en la condición de *Complejidad aritmética* 1 (68% aciertos) que en la de *Complejidad aritmética* 2 (24% aciertos) y que en la de *Complejidad aritmética* 4 (0% aciertos), siguiendo los resultados, por lo tanto, una gradación muy clara relacionada con la complejidad aritmética.

En lo referente a las comparaciones entre la tarea clásica usada por Gigerenzer y la tarea de probabilidad en sus distintos niveles de complejidad, los resultados fueron en línea con lo anterior, es decir, dependientes de complejidad. La primera comparación correspondería con la clásica realizada entre probabilidades y frecuencias naturales en el experimento de Gigerenzer y Hoffrage (1995), y es también equivalente a una comparación entre lo que en el experimento anterior hemos llamado probabilidades relativas y frecuencias absolutas. Comparamos nuestra condición de frecuencias con una *Complejidad aritmética* = 2 (23% aciertos) con la condición de probabilidades con una *Complejidad aritmética* = 4 (0% aciertos). La diferencia entre ellas es, tal y como sugieren tantos otros, notable y significativa, t’(30)= 2,958, p= 0,006.

Lo interesante viene al reducir la complejidad aritmética de las condiciones de probabilidades. Al comparar frecuencias (*Complejidad aritmética* = 2) (23% aciertos) con probabilidades (*Complejidad aritmética* = 2) (26% aciertos), es decir, probabilidades y frecuencias con complejidad aritmética igualada, la diferencia entre ellas desaparece, t(60)= -0,292, p= 0,771. Finalmente, cuando comparamos frecuencias (*Complejidad aritmética* = 2) (23% aciertos) con probabilidades (*Complejidad aritmética* = 1) (68% aciertos), donde, por primera vez, la complejidad aritmética en probabilidades es menor que en frecuencias, encontramos una diferencia muy amplia y significativa entre ambas, t(60)= -3,944, p< 0,001, pero a favor de las probabilidades o, dicho de otra manera, a favor de la complejidad baja.

Tabla 20. Porcentaje de respuestas correctas de las respuestas de los participantes para cada condición experimental.

**Aciertos Intervalo de Confianza (95%) N**

**Probabilidades**

**1** 68 53,9 / 81,6 31

**2** 26 12,0 / 39,6 31

**4** 0 -13,8 / 13,8 31

**Frecuencias**

**2** 23 8,8 / 36,4 31

En este experimento hemos disociado, en la medida de lo posible, los formatos de representación y la complejidad aritmética. Encontramos un efecto muy importante de complejidad y aportamos pruebas congruentes con que las diferencias entre probabilidades y frecuencias pueden ser explicadas de forma clara a partir de diferencias en complejidad aritmética. Que nosotros sepamos, ésta es la primera vez que se demuestra empíricamente la influencia real aislada de la complejidad aritmética, más allá de las diferencias entre formatos de representación y otros factores.

### 2.4. Complejidad aritmética y sus matices

En este último experimento pretendemos poner a prueba una de las afirmaciones que realizan Gigerenzer y Hoffrage (1995) y que resulta fundamental para diferenciar nuestra postura de la de ellos, además de afianzar el principio de pasos de cálculo.

En el experimento anterior comprobamos que la complejidad computacional o aritmética, definida como el número de operaciones mínimas necesarias para resolver el problema, es una buena medida de dificultad y que es extraordinariamente precisa a la hora de predecir el rendimiento en problemas de cálculo de probabilidad Bayesiana. Pero como hemos visto, una de las diferencias que ya indican los frecuentistas entre probabilidades y frecuencias naturales es la complejidad computacional. Ésta es definida por ellos a partir de los algoritmos Bayesianos, pero es usada casi de forma exclusiva para distinguir entre probabilidades y frecuencias, y cuando no, se usa de manera burda y contradictoria.

Tanto es así, que Gigerenzer y Hoffrage (1995) realizan una predicción (predicción 3) que se refiere a la independencia de los resultados en función del menú de la información. Esta predicción contradice frontalmente el principio de complejidad computacional que ellos mismos propugnan.

Uno de los ejes principales defendidos por los frecuentistas es que los cálculos Bayesianos son más sencillos cuando se trabaja con frecuencias naturales. Gigerenzer y Hoffrage distinguen entre dos aspectos de la representación de la información, el formato de la información (probabilidades o frecuencias) y el menú de la información (estándar y abreviado). Del formato de la información ya hemos hablado, pero ¿qué es el menú de la información? Veamos cómo son exactamente los distintos formatos y menús.

Tabla 21. Formatos y menús de la información en el problema clásico de mamografías (Gigerenzer y Hoffrage, 1995).

|  |
| --- |
| **Menú estándar de probabilidades** |
|  |
| *La probabilidad de tener cáncer de mama es del 1% para mujeres de cuarenta años que participan en chequeos rutinarios.*  *Si una mujer tiene cáncer de mama, la probabilidad de que una mamografía le dé positivo es del 80%.*  *Si una mujer no tiene cáncer de mama, la probabilidad de que una mamografía le dé también positivo es del 9,6%.*  *A una mujer de este grupo de edad le dio positivo una mamografía en un chequeo rutinario. ¿Cuál es la probabilidad de que tenga realmente cáncer de mama?\_\_\_\_\_\_%* |
|  |
| **Menú estándar de frecuencias** |
|  |
| *10 de cada 1000 mujeres de cuarenta años que participan en chequeos rutinarios tienen cáncer de mama.*  *8 de cada 1000 mujeres de cuarenta años que participan en chequeos rutinarios tienen cáncer de mama y una mamografía positiva.*  *95 de cada 1000 mujeres de cuarenta años que participan en chequeos rutinarios no tienen cáncer de mama, pero la mamografía es positiva.*  ***De un nuevo grupo de mujeres de cuarenta años a las que les dio positivo una mamografía en un chequeo rutinario, ¿cuántas esperas que tengan realmente cáncer de mama?*** *\_\_\_\_\_ de \_\_\_\_\_* |
|  |
| **Menú abreviado de probabilidades** |
|  |
| *La probabilidad de que a una mujer de cuarenta años le dé positivo una mamografía en un chequeo rutinario es del 10,3%.*  *La probabilidad de tener cáncer de mama y una mamografía positiva es del 0,8% para una mujer de cuarenta años que participa en un chequeo rutinario.*  *A una mujer de este grupo de edad le dio positivo una mamografía en un chequeo rutinario. ¿Cuál es la probabilidad de que tenga realmente cáncer de mama?\_\_\_\_\_\_%* |
|  |
| **Menú abreviado de frecuencias** |
|  |
| *A 103 de cada 1000 mujeres de cuarenta años les da positivo una mamografía en un chequeo rutinario.*  *8 de cada 1000 mujeres de cuarenta años que participan en chequeos rutinarios tienen cáncer de mama y una mamografía positiva.*  ***De un nuevo grupo de mujeres de cuarenta años a las que les dio positivo una mamografía en un chequeo rutinario, ¿cuántas esperas que tengan realmente cáncer de mama?*** *\_\_\_\_\_ de \_\_\_\_\_* |

Ellos lo definen como la forma en la que la información está segmentada en cualquiera de los formatos. Según Gigerenzer y Hoffrage, el menú estándar de probabilidades muestra tres piezas, la probabilidad previa (*base-rate*), el ratio de aciertos (*hit rate*) y el ratio de falsas alarmas (*false alarm rate*) (p(H), p(D|H) y p(D|~H), respectivamente, mientras que con frecuencias naturales es suficiente con dos, D&H y D&~H, o alternativamente D&H y D, en lo que llaman menú abreviado. A partir de aquí, se pueden combinar ambas variables creando versiones estándar de frecuencias y cortas de probabilidades. La diferencia de complejidad computacional entre unos y otros deriva en los resultados teóricos 5 y 6 (Gigerenzer y Hoffrage, 1995, pag. 688), que indican que habrá diferencias en probabilidades, pero no frecuencias al comparar el menú estándar con el abreviado.

Y, ¿cómo se resuelven los problemas en la versión abreviada? En el caso de los problemas de probabilidades con el menú abreviado la fórmula es:

**(1)** 

Y con problemas de frecuencias con el menú abreviado:

**(2)** 

Para Gigerenzer y Hoffrage, con frecuencias naturales siempre se usa el mismo algoritmo **(2)** para resolver los problemas tanto en su forma estándar como abreviada, al contrario que con probabilidades, donde la versión estándar es más compleja computacionalmente que la abreviada:

**(3)** 

Así pues, a modo aclaratorio, en el cuadro siguiente se pueden ver las distintas condiciones y los algoritmos necesarios para su resolución:

Cuadro 15. Condiciones y los algoritmos necesarios para su resolución planteados por (Gigerenzer y Hoffrage, 1995).

**Menú Estándar**

**Probabilidades**

**(3)** 

**Frecuencias naturales**

**(2)** 

**Menú Abreviado**

**Probabilidades**

**(1)** 

**Frecuencias naturales**

**\*(1)** 

\*Gigerenzer y Hoffrage indican que el algoritmo necesario para resolver el menú abreviado es **(2)** o **(1)** indistintamente, pero en su material la forma usada corresponde a **(1)**.

Nos interesa especialmente la inicial equivalencia que Gigerenzer y Hoffrage (1995) establecen entre los algoritmos **(1)** y **(2)**,accidentalmente mostrada a partir de la diferencia entre menús estándar y abreviado para frecuencias. Ellos parten de la hipótesis de que la proporción de respuestas correctas (Bayesianas) para el formato frecuentista es independiente de menú y, por lo tanto, que, como decíamos, **(1)** y **(2)**,son virtualmente idénticos.Al ofrecer los resultados y no desaparecer completamente el efecto de menú matizan que en el caso de **(2)** hay un paso más de cálculo que podría explicar la *pequeña* diferencia de 3,5 puntos encontrada entre **(1)** y **(2)** a favor del primero.

Parece, por tanto, que, a pesar de que Gigerenzer y Hoffrage introducen teóricamente que la complejidad computacional tiene un papel relevante en la discusión que nos ocupa, intentan obviar las ramificaciones de ésta para centrarse en las diferencias existentes entre probabilidades y frecuencias naturales.

Aquí vamos a centrarnos en esas ramificaciones comparando dos condiciones con frecuencias naturales con distintos pasos de cálculo, condiciones entre las que según Gigerenzer y Hoffrage, teóricamente no debería haber apenas diferencias.

Cuadro 16. Complejidad (en pasos de cálculo necesarios) de las distintas condiciones del experimento 7.

**Complejidad aritmética**

**Pasos de cálculo necesarios: 1**



**Pasos de cálculo necesarios: 2**



#### Método

#### Participantes y diseño

Un total de 154 estudiantes de psicología de la Universidad de La Laguna completaron un cuestionario en clase. Fueron asignados de forma aleatoria a alguno de los dos grupos correspondientes al factor entre-sujetos *Complejidad aritmética (pasos de cálculo necesarios)* (1, 2).

Tabla 22. Variable usada y representación esquemática del material.

**Variable independiente**

Entre-sujetos:

* Complejidad aritmética
  + 1 paso de cálculo
  + 2 pasos de cálculo

#### Procedimiento y materiales

Los participantes contestaron a una sola pregunta relativa a la probabilidad de que una mujer tuviera cáncer de mama dado que el resultado de la mamografía había dado positivo.

En la tabla siguiente se puede ver el material experimental:

Tabla 23. Material del experimento 7 (en cursiva).

***Instrucciones***

*Intenta resolver el siguiente problema de razonamiento sobre mamografías. La mamografía es una prueba médica que nos permite detectar pequeños bultos en las mamas (un resultado positivo indica la presencia de bultos que pueden ser o no cancerígenos). Este método se usa para ayudar a diagnosticar de forma temprana el cáncer de mama.*

*Lee la información del problema atentamente y escribe tu respuesta. Puedes tomar todo el tiempo que quieras para resolver el problema.*

*Se realiza un estudio a partir de datos de un gran número de mujeres.*

* 1 paso de cálculo

*A 103 de cada 1000 mujeres de cuarenta años les da positivo una mamografía en un chequeo rutinario.*

*8 de cada 1000 mujeres de cuarenta años que participan en chequeos rutinarios tienen cáncer de mama y una mamografía positiva.*

* 2 pasos de cálculo

*8 de cada 1000 mujeres de cuarenta años que participan en chequeos rutinarios tienen cáncer de mama y una mamografía positiva.*

*95 de cada 1000 mujeres de cuarenta años que participan en chequeos rutinarios no tienen cáncer de mama, pero la mamografía es positiva.*

***De un nuevo grupo de mujeres de cuarenta años a las que les dio positivo una mamografía en un chequeo rutinario, ¿cuántas esperas que tengan realmente cáncer de mama?*** *\_\_\_\_\_ de \_\_\_\_\_*

A continuación, en la tabla 24, se puede ver una representación esquemática de las distintas condiciones experimentales, así como de la manera de resolverlas.

Tabla 24. Representación esquemática del material del experimento y pasos de cálculo necesarios para su resolución.

**Complejidad aritmética = 1**

Pasos de cálculo necesarios: 1



**Complejidad aritmética = 2**

Pasos de cálculo necesarios: 2



#### Resultados

Los resultados se presentan en la tabla 25. Comparábamos las condiciones en formato de frecuencias naturales con complejidad aritmética 2 y complejidad aritmética 1. Recordamos que los Gigerenzer y Hoffrage (1995) dicen explícitamente que ambas condiciones son equivalentes y no encuentran diferencias entre ellas.

El ANOVA mostró un efecto significativo de la *Complejidad aritmética* (F(1, 153)= 16,085, MSE= 0,150, p< 0,0001, η2=0,09), respondiendo mejor los participantes en la condición de *Complejidad aritmética* 1 (31% aciertos) que en la de *Complejidad aritmética* 2 (7% aciertos), por lo que los resultados mostraron una diferenciación muy clara relacionada con la complejidad aritmética.

Tabla 25. Porcentaje de respuestas correctas de las respuestas de los participantes para cada condición experimental.

**Aciertos N**

**Complejidad aritmética**

**1** 31 78

**2** 7 76

Estos resultados suponen el último clavo en el ataúd frecuentista, destacando de manera más clara aún, si cabe, la importancia del número de pasos de cálculo sobre el rendimiento en los problemas de cálculo Bayesiano, independientemente del formato de representación, y contradiciendo de manera directa una de las predicciones realizadas por Gigerenzer y Hoffrage (1995) en referencia a la independencia de los resultados en función del menú de la información.

### Discusión: Complejidad y formatos de representación con probabilidades Bayesianas

Diferentes factores pueden ser y han sido usados para demostrar cómo la complejidad computacional afecta al rendimiento de las personas (e.g. *nested sets* (Sloman, Over, Slovak, y Stibel, 2003; Tversky y Kahneman, 1983). Por desgracia, el solapamiento entre predicciones, algunos malentendidos y sutilezas sobre la naturaleza de las frecuencias naturales han hecho que estos esfuerzos no resulten concluyentes. La única manera de zanjar finalmente esta disputa es usando problemas donde, controlando la complejidad, se varíe el formato de representación.

A lo largo de los diferentes experimentos de este segundo bloque hemos tratado de limar progresivamente la hipótesis presentada por el campo frecuentista, referente a la menor dificultad que entrañan los problemas Bayesianos al usar frecuencias naturales para su representación. Entre las ventajas generalmente aceptadas de las frecuencias naturales se cuentan la conservación de las probabilidades previas y una estructura más clara (Brase y Barbey, 2006; Gigerenzer y Hoffrage, 1995; Hoffrage et al., 2002; Macchi, 2000). Son estas ventajas y la presencia de las frecuencias a lo largo de nuestra historia evolutiva las que, según los frecuentistas, han derivado en una especialización cerebral para el tratamiento de información probabilística en forma de frecuencias naturales.

Como ya hemos comentado, los argumentos evolutivos pueden resultar enormemente interesantes, pero difíciles de afrontar empíricamente. Es por ello que, como decíamos, nos hemos centrado en la primera dimensión presentada por los frecuentistas para justificar la superioridad de las frecuencias naturales, esto es, la menor complejidad de resolver problemas Bayesianos con éstas. Hasta donde nosotros sabemos, esta afirmación se basa únicamente en argumentos teóricos extraídos a partir de la manera de resolver los problemas en uno y otro formato (ya que en los resultados empíricos no se controla la complejidad computacional). Hemos intentado falsar este supuesto en los distintos experimentos, buscando situaciones en las que despojar a las frecuencias naturales del estatus especial que ostentan. Situaciones en las que la complejidad algorítmica de frecuencias y probabilidades fuera comparable. Hemos refinado nuestra crítica en un recorrido que ha partido de pruebas tentativas a partir de la manipulación de las probabilidades previas, pasando por una comparación con unos cimientos teóricos más fuertes basados en una dimensión con mayor poder descriptivo, como es la clase referencia de la información (Absoluta/Relativa) hasta llegar a un nivel explicativo que consideramos nuclear, como es la complejidad aritmética (número de pasos de cálculo), ofreciéndonos este último constructo una capacidad predictiva extraordinaria. Esto nos permitió no solo conseguir resultados en los que dejáramos a un lado la complejidad de uno u otro formato de representación y al mismo tiempo las diferencias de rendimiento de los sujetos, sino que llegamos a encontrar situaciones en las que la menor dificultad de problemas clásicos Bayesianos en formato probabilístico daba lugar a unos resultados muy por encima de los de los formatos frecuentistas. Por último, la aplicación consistente de este principio nos permitió refinar y aclarar las predicciones frecuentistas con respecto a diferencias de dificultad entre menús de información, dejando claro que el uso que han hecho los frecuentistas del principio de complejidad computacional es una mera excusa para diferenciar entre probabilidades y frecuencias naturales.

A pesar de su dificultad, también nos sumergimos levemente en los aspectos evolutivos, usando para ello pruebas cognitivas. La diferencia encontrada en las tareas Bayesianas clásicas para los grupos superior e inferior del CRT va en contra de la afirmación frecuentista sobre la existencia de un algoritmo o módulo innato. La existencia de correlaciones entre inteligencia general y resultados en tareas Bayesianas no tiene sentido si se tratara de un algoritmo especializado e independiente (De Neys, 2007).

Estos resultados nos permiten finalmente establecer un acuerdo de mínimos sobre el que construir posteriores discusiones. No existe ambigüedad ni apenas supuestos en nuestra operacionalización empírica. Usamos la complejidad aritmética definida como el número de pasos de cálculo necesarios para poder resolver una tarea de modo similar a lo planteado por Gigerenzer (Gigerenzer y Hoffrage, 1995), solo que de manera consistente, es decir, no con la intención de señalar las deficiencias de un formato o condición particular, sino para definir de manera más precisa las diferentes condiciones usadas históricamente en el campo. Esto nos permite explicar diferencias que han sido dejadas de lado habitualmente y que ayudan a cerrar disputas basadas en conceptos de alto nivel construidos sobre falsos supuestos.

En lo que respecta a la segunda dimensión frecuentista, la existencia de un módulo o algoritmo especializado para las frecuencias naturales, hemos visto no solo que hay variables que nos permiten explicar mejor las diferencias (complejidad aritmética) y, por lo tanto, cuestionan la teórica superioridad de las frecuencias naturales, sino que, además, existen correlaciones entre pruebas cognitivas (CRT) y aciertos en las tareas clásicas Bayesianas, lo que sería, al menos, poco probable en caso de existir un módulo o algoritmo especializado, como proponen algunos autores (Cosmides y Tooby, 1996; Gigerenzer y Hoffrage, 1995).

*“Get your facts first, and then you can distort them as much as you please.”*

**Mark Twain**

# DISCUSIÓN GENERAL

### A vista de pájaro

En este trabajo hemos pretendido crear un marco común que nos permita entender la extensa literatura y las contradicciones existentes en el campo del cálculo de probabilidad. Empezamos investigando, en la primera parte, si la superioridad clásica de las frecuencias naturales sobre las probabilidades se daba también en problemas de probabilidad simple. Quisimos averiguar cuál era la influencia real de varios factores contextuales: cosas como la presencia explícita de la clase complementaria (pimientos no picantes al preguntar sobre pimientos picantes) o la coincidencia en escala entre instrucciones y respuesta. Los resultados mostraron que, a idéntica complejidad computacional, se obtienen iguales resultados. No encontramos diferencias entre probabilidades y frecuencias naturales con un paradigma de cálculo de probabilidades simples, tampoco en función de la presencia o ausencia de la clase complementaria. Por otro lado, sí vimos cómo los sujetos eran mejores al haber coincidencia, en escala, entre las instrucciones y la respuesta, es decir, al evitar un paso en el cálculo, o, dicho de otro modo, al no haber correspondencia entre la representación de la situación y la descripción del plato, lo que evitaba que incurrieran en la *falacia del jugador* en mayor medida.

Mientras la aproximación frecuentista no puede explicar estos resultados, especialmente los relativos a la influencia de una variable extensional como son las posibilidades representadas, otras teorías sí pueden. En particular, aquellas que sostienen que las frecuencias naturales facilitan el razonamiento Bayesiano porque habitualmente dan pistas extensionales sobre el grupo de inclusión (Johnson-Laird et al., 1999; Kahneman & Tversky, 1996; Sloman, Over, & Slovak, 2003). Un rasgo fundamental de estas teorías es que no asumen racionalidad como propiedad central de sus mecanismos. Por lo tanto, los mismos mecanismos de los que surge el pensamiento racional pueden, en determinadas circunstancias, provocar errores o sesgos.

En la segunda parte quisimos comprobar si el principio surgido de la primera parte, el que la complejidad computacional podría explicar en mayor medida los resultados que cualquier otro factor, se podía extender al cálculo Bayesiano. Y adaptamos los problemas clásicos de probabilidades y frecuencias naturales para intentar igualar la complejidad. Empezamos por igualar las probabilidades previas, para continuar con retoques de clases de referencia, con manipulaciones más y más finas de la complejidad aritmética, entendida como los pasos de cálculo necesarios. Aunque también tuvimos tiempo para comprobar hasta qué punto habilidades cognitivas generales correlacionaban con el acierto de los sujetos en problemas Bayesianos. Lo que descubrimos había ya sido comentado en parte por el frente frecuentista y es que, efectivamente, la complejidad de los problemas es importante. El matiz que añadimos es que esto no es así en el sentido de que las frecuencias naturales sean más sencillas necesariamente. Resulta que al igualar probabilidades previas (y, por lo tanto, reducir disparidades en el cálculo) desaparecen las diferencias entre probabilidades y frecuencias naturales. Al controlar los pasos de cálculo necesarios podemos crear situaciones en las que las probabilidades sean tanto o más sencillas que las frecuencias naturales, y viceversa. Y no sólo eso, sino que también podemos explicar los resultados de la literatura y predecir con gran exactitud nuevos resultados, además de aclarar algunas confusiones del campo frecuentista, como el papel que representan los menús de la información.

Por otro lado, vimos cómo una prueba que podría ser considerada de habilidades cognitivas generales, como el CRT, correlaciona largamente con los resultados en las condiciones clásicas de los problemas Bayesianos, y no solo eso, sino que la correlación es incluso mayor allí donde debería ser menor si realmente existiera un módulo o algoritmo cognitivo especializado en frecuencias naturales (en la condición de frecuencias).

En los resultados de esta segunda parte, las predicciones a partir del número de pasos de cálculo no se diferencian de aquellas realizadas por teorías como la de los Modelos Mentales (Johnson-Laird et al., 1999). De hecho, el *principio del subconjunto* (*subset principle*) – asumiendo equiprobabilidad, una probabilidad condicionada p(H|D) depende del subconjunto de D que son H – propuesto por Johnson-Laird (1999) como algoritmo usado para resolver problemas Bayesianos, corresponde a la fórmula (2) del trabajo de Gigerenzer (1995, pag. 687). Ésta es, a su vez, similar a la fórmula que vemos a continuación:



Las operaciones de cálculo predichas por los Modelos Mentales a partir del *principio del subconjunto* unido al número de posibilidades tenidas en cuenta, no se diferencian de las propuestas por Gigerenzer (1995) para los problemas que se usan habitualmente en la literatura, y en este trabajo en particular. Las operaciones necesarias para resolver varios problemas, así como los modelos mentales representados, se pueden ver en la siguiente tabla.

Tabla 26: Modelos mentales para las distintas condiciones experimentales. En negrita, los modelos necesarios para la resolución y, entre paréntesis, la operación usada.

**Probabilidades – Complejidad aritmética = 0/1**

Probabilidad P. Condicional Modelos

**Enfermedad 10% Síntoma 100% p(H)**

~Síntoma 0%

~Enfermedad 90% Síntoma 0%

~Síntoma 100%



**Probabilidades – Complejidad aritmética = 2**

Probabilidad P. Condicional Modelos

**Enfermedad 10% Síntoma 100% p(H)**

~Síntoma 0%

**~Enfermedad 90% Síntoma 100% p(~H)**

~Síntoma 0%



**Probabilidades – Complejidad aritmética = 4**

Probabilidad P. Condicional Modelos

**Enfermedad 10% Síntoma 80% p(H), p(D|H)**

~Síntoma 20%

**~Enfermedad 90% Síntoma 10% p(~H), p(D|~H)**

~Síntoma 90%



**Frecuencias Naturales – Complejidad aritmética = 2**

Casos Modelos

**Enfermedad Síntoma 80 p(H&D)**

Enfermedad ~Síntoma 20

**~Enfermedad Síntoma 90 p(~H&D)**

~Enfermedad ~Síntoma 810



Alternativamente, en la condición de complejidad aritmética 4, que tiene cuatro modelos, se pueden combinar los pares de modelos p(H) y p(D|H) por un lado y p(~H) y p(D|~H) por otro, formando p(D&H) y p(D&~H), respectivamente. Esto reduciría a dos el número de modelos, pero con la contrapartida de exigir una computación extra, en la línea del principio de complejidad computacional descrito en este trabajo.

El número de modelos mentales coincide con el número de pasos de cálculo necesarios – extraídos a partir de los algoritmos Bayesianos – para las situaciones mostradas, lo que nos impide distinguir entre ambas explicaciones, como decíamos antes.

Aunque, la teoría de los modelos mentales haría una predicción adicional. Basándose en el principio de verdad, que dice que solo se crean representaciones de lo que es cierto (Johnson-Laird et al., 1999, pag. 68), se predeciría que el tipo de error más común debería ser el que consistiera en desestimar el dato de los que no tienen enfermedad o síntoma, lo que es congruente con los datos presentados por Gigerenzer y Hoffrage (1995, pag. 695). En el futuro nos gustaría continuar esta línea de investigación, profundizando en la distinción entre pasos de cálculo y modelos mentales en probabilidades condicionadas.

Así pues, el conjunto de los resultados señala hacia la complejidad aritmética como el factor más sencillo y con más potencia explicativa, y no solo para cálculo de probabilidad Bayesiano, también para el simple. En los aspectos evolutivos, los resultados son tal vez menos directos, pero igualmente contundentes.

### Sobre pálidos lagartos

Es crucial para la supervivencia del ser humano como especie el adecuado ajuste al entorno. Desarrollamos, como buenos mamíferos, sistemas de termorregulación que nos permitieron ampliar nuestros horizontes de una manera que haría palidecer de envidia a los reptiles, si esto fuera remotamente posible, e innumerables otras adaptaciones al medio físico. Pero no todas las herramientas para la supervivencia que aparecieron fueron de naturaleza física, en su sentido menos extenso (piernas, pulgares, visión estereoscópica, etc.), metabólica (respiración, digestión, termorregulación, etc.), o similares. Posiblemente, el instrumento más potente de cara a nuestra supervivencia que desarrollamos fuera la hipertrofiada masa de neuronas a la que llamamos cerebro.

Dentro de la psicología cognitiva nos interesa estudiar cómo funciona exactamente éste, para lo que puede resultar de ayuda conocer cómo está estructurado. Se trata de una herramienta de procesamiento de información de carácter general, una colección de módulos o sistemas relativamente indiferenciados o, por el contrario, sumamente compartimentados e impermeables. En este trabajo no vamos a dar respuestas de esta magnitud, pero nos adentraremos en un debate que bebe de la polémica sobre cómo funciona la mente.

Algunas de las ventajas que aporta el cerebro de cara a la supervivencia son una capacidad de simulación, predicción, extrapolación, toma de decisiones, deducción, inducción, y un larguísimo etcétera, sin precedentes. Se podría decir sin miedo a equivocarnos que al menos una de las bases de estas capacidades es el cálculo de probabilidad. En concreto, el cálculo de probabilidades condicionadas guía nuestra habilidad para inducir, para entender cómo funciona el mundo. No somos la única especie que cuenta con unas habilidades casi mágicas en este sentido. Si bajamos en la escala evolutiva tratando de encontrar pruebas de las asombrosas proezas Bayesianas de las que son capaces los humanos podemos llegar muy, muy abajo. Gigerenzer nos recuerda que en investigaciones sobre caza y recolección animal con abejas, patos, ratas y hasta hormigas, se comprobó que éstas se comportan como si fueran buenos estadísticos intuitivos, sensibles a cambios en frecuencias en el entorno (Gallistel, 1990; Real, 1991; Real & Caraco, 1986).

Nos interesa, por lo tanto, definir cómo funciona exactamente esto, cómo los humanos calculan probabilidades y qué factores determinan su éxito o fracaso. Vimos que a la hora de realizar cálculos Bayesianos hay una extensa literatura mostrando que el formato de representación es determinante. Los frecuentistas basan, como hemos comentado con anterioridad, la aparente superioridad de las frecuencias en dos ejes:

1) Los algoritmos Bayesianos son más sencillos computacionalmente cuando la información está codificada en forma de frecuencias naturales (Gigerenzer y Hoffrage, 1995).

2) Asumiendo que los humanos han desarrollado a través de la evolución algoritmos cognitivos, éstos habrán sido diseñados para trabajar con frecuencias naturales (Cosmides y Tooby, 1996; Gigerenzer, 1991, 1993; Gigerenzer y Hoffrage, 1995).

Vamos a tratar de encajar el conjunto de resultados presentado en este trabajo de tal manera que demuestren que, a un nivel determinado, sólo es necesaria la complejidad computacional para explicar el rendimiento de los humanos al enfrentarse al cálculo de probabilidad Bayesiano y, además, que los argumentos evolutivos son un hermoso añadido carente de sentido. También pretendemos destruir el mito de que las frecuencias naturales son intrínsecamente más sencillas.

### Complejidad Aritmética

Como decíamos antes, la complejidad aritmética (número de pasos de cálculo) nos ofrece una capacidad predictiva extraordinaria. Esto nos permitió, como pudimos observar en el experimento 2.3., controlar, a partir de manipulaciones mínimas, el nivel de dificultad y, por ende, el rendimiento de los sujetos en problemas Bayesianos clásicos.

Pero, hasta cierto punto, esta afirmación no es nueva. Vimos cómo en el primero de los ejes en los que se basan para declarar la superioridad de las frecuencias naturales ya se estipula que el formato frecuentista hace más sencilla la resolución de problemas Bayesianos. Esto es así por varios motivos, entre ellos, la conservación de probabilidades previas y la menor complejidad algorítmica en la resolución de los problemas. Hemos visto, no obstante, cómo la conservación de probabilidades previas no es exclusiva de las frecuencias naturales. Las probabilidades con clase de referencia absoluta p(D&H), a diferencia de las probabilidades con clase de referencia relativa p(D|H), también tienen esta propiedad.

De todos modos, el compromiso que los frecuentistas establecen con el principio de complejidad aritmética es más bien parco. Lo usan para diferenciar entre probabilidades y frecuencias naturales, pero no exploran las consecuencias de este principio. Lo vimos en la igualdad que propugnan entre los siguientes algoritmos:

**(1)**  **(2)** 

Como dijimos antes, ellos parten de la hipótesis de que la proporción de respuestas correctas (Bayesianas) en **(1)** y **(2)** debería ser virtualmente idéntica.Sus resultados contradicen de manera leve esto, obligándoles a matizar que en el caso de **(2)** hay un paso más de cálculo que podría explicar la *pequeña* diferencia de 3,5 puntos encontrada entre **(1)** y **(2)** a favor del primero.

Nuestros resultados, que suponen una comparación directa de ambos algoritmos, no pueden ser más claros. Encontramos una diferencia superior a los 20 puntos en el porcentaje medio de aciertos entre ambas condiciones, lo que apoya de manera rotunda la importancia que tienen los pasos de cálculo por encima de distinciones relativas al formato o al menú de la información.

Cuando Gigerenzer y Hoffrage (1995) plantean los problemas manipulando el formato y el menú de la información, realizan una serie de elecciones arbitrarias sobre qué datos aportar o no, obviando las consecuencias a nivel de complejidad computacional. Esto se demuestra en la igualdad que establecen, y que hemos visto es falsa, entre los algoritmos **(1)** y **(2)**.

Nuestros resultados muestran que es necesario no dejar de lado en ningún momento los fundamentos de la complejidad computacional, que modificando la información que se aporta sobre los distintos elementos de la ecuación (tanto por omisión como por el efecto que una cifra distinta pueda tener sobre el cálculo, o el de la clase de referencia, o de la pregunta) es posible crear condiciones arbitrariamente complicadas o sencillas independientemente del formato de la información.

Por ejemplo, imaginemos que cogemos los problemas clásicos con probabilidades relativas y frecuencias absolutas y, en lugar de dar la información sobre la cantidad o porcentaje de mamografías positivas y negativas relativa a las mujeres con cáncer o no de mama, invertimos la estructura y damos información sobre mujeres con cáncer de mama o no relativa a las mujeres con mamografía positiva o negativa. Dicho de modo más sencillo, en lugar de hablar de el número de mujeres con cáncer a las que les ha dado positivo la mamografía, hablamos de el número de mamografías positivas que corresponden a mujeres que tienen cáncer de mama. En este caso, la estructura computa por nosotros la respuesta, p(H|D), para las probabilidades relativas y, sin embargo, aún es necesario realizar una operación en el caso de las frecuencias absolutas. Es decir, como se puede ver en la tabla siguiente, p(Cáncer|Mamografía+) (7.8%, la respuesta correcta) es uno de los datos que se aporta en el caso de las probabilidades relativas al invertir la estructura. Sin embargo, con frecuencias absolutas, la información que nos da el problema nos obliga a coger dos piezas de información (8/103). Y no solo eso, dependiendo de qué información en concreto aportemos, o para usar la terminología de Gigerenzer y Hoffrage (1995), del menú de la información, podemos aumentar el número de pasos de cálculo necesarios a dos (8/(8+95) en el caso de las frecuencias naturales, simplemente no mostrando el número de mamografías positivas).

Si alguien elaborara su estudio usando esta estructura, las conclusiones podrían ser que las probabilidades de eventos simples simplifican la computación y, por lo tanto, son superiores a las frecuencias naturales. Si esta persona fuera, además, especialmente avezada, podría decir que, puesto que hay un sistema innato de número que se basa en la comparación bruta de proporciones (i.e. probabilidades; S. Dehaene, 1997; S Dehaene, 2001; Feigenson et al., 2004; Spelke y Kinzler, 2007), la selección natural parece haber favorecido a las probabilidades.

Podemos ver esta manipulación y la variación de la complejidad aritmética que supone en la siguiente tabla.

Tabla 27. Representación de las condiciones clásicas e invertidas (en azul claro, la información que no se aporta en el problema).

**Probabilidades**

Pasos de cálculo: 4/5



**Frecuencias**

Pasos de cálculo: 2



**Probabilidades - Invertidas**

Pasos de cálculo: 0



**Frecuencias – Invertidas**

Pasos de cálculo: 2(\*)



(\*) Si en la información del problema se aporta también p(D) entonces los pasos de cálculo pasan a ser 1 (8/103).

No es una de nuestras pretensiones defender la superioridad de las probabilidades sobre las frecuencias ni nada remotamente similar. El ejemplo anterior únicamente demuestra que, seleccionando fragmentos arbitrarios de la realidad y proponiéndolos como relevantes, se pueden llegar a conclusiones igualmente arbitrarias. En cualquier caso, no creemos que la hipótesis frecuentista haya nacido a partir de una selección consciente y malintencionada de datos, pero sí con poco calado teórico, o al menos, con un sesgo claro.

Hay una limitación en las posibilidades de manipulación de las que hablamos y es que, es cierto que las frecuencias naturales conservan información sobre probabilidades previas, lo que de manera automática supone una ventaja sobre las probabilidades relativas en muchos casos, aunque no sobre probabilidades con una clase de referencia absoluta, entre otras. Si nos dan una frecuencia natural, por ejemplo, 8 de cada 1000 mujeres con mamografías positivas tienen cáncer de mama, ese 8 de cada 1000 es en realidad D&H (Mamografía+ & Cáncer). El dato equivalente con probabilidades relativas sería p(D|H), la probabilidad de una mamografía positiva dado que una mujer tiene cáncer, 80%, a lo que habría que añadir la probabilidad previa, esto es, la probabilidad de tener cáncer p(H), o un 1%. Como vemos, por lo tanto, en un caso tenemos directamente la probabilidad previa integrada en D&H (8 de cada 1000) y en el otro la tenemos que añadir p(D|H), p(D|H)\*p(H) (80%\*1%). De todos modos, esta limitación es parcialmente ficticia. Ya antes mostramos cómo con probabilidades absolutas las cosas cambian. Con éstas se aporta directamente p(D&H), 0,8%, de igual modo que con frecuencias naturales.

En cualquier caso, a partir de este tipo de análisis queda claro que la distinción que realizamos en el experimento 2.2. entre clase de referencia absoluta y relativa es esencial, incluso más allá que la diferencia entre probabilidades y frecuencias, al menos, en lo que respecta a la conservación de probabilidades previas.

Cuadro 17. Clase de referencia absoluta y relativa

**Clase de referencia Absoluta**

*Conservan información sobre probabilidades previas*

**Probabilidades:** *p(D&H) = 0,8%*

**Frecuencias Naturales:** *(D&H) = 8 de cada 1000*

**Clase de referencia Relativa**

*No conservan información sobre probabilidades previas*

**Probabilidades:**  *p(D|H) =80%*

**Frecuencias Naturales (\*):** *(D|H) = 8 de cada 10*

**(\*)** En este caso, es discutible que no conserve probabilidades previas.

El principio fundamental que pretendemos transmitir es que en la disputa sobre formatos de representación de la información y cálculo Bayesiano se han obviado explicaciones más parsimoniosas e inclusivas a cambio de grandilocuentes teorías basadas en información parcial.

Pero los defensores de la *Ecological Rationality* no hablan únicamente de complejidad, insisten en la existencia de un módulo o algoritmo cognitivo que subyace a estas diferencias. En la discusión subsiguiente presente en la literatura, los aspectos evolutivos tienen un peso enorme y son los que más reactividad causan. Hay también algunos intentos por despojar a las frecuencias naturales, no solo de su estatus cuasi mítico de formato de representación favorecido por la evolución, sino también del aura de necesidad que la teórica menor complejidad computacional de las frecuencias tiene.

Nuestra intención con este trabajo, al menos una de ellas, es devolver la discusión al cauce que creemos nunca debiera haber abandonado, la complejidad computacional, y cómo ésta afecta al cálculo. El formato de representación es un nivel de análisis demasiado grueso, como intentar pintar una miniatura con brocha, y su uso hace que se pierdan los matices. Matices como la clase de referencia de la información, que es vital para entender cómo los humanos se enfrentan al cálculo de probabilidad Bayesiano y que se puede entender mejor usando el pincel de la complejidad computacional o algorítmica.

Como hemos visto en nuestros resultados, las frecuencias naturales son mejores que las probabilidades cuando se dan una serie de condiciones, no por defecto, y es fácil crear un set de condiciones que haga desaparecer las diferencias o incluso las invierta.

Es importante destacar que la explicación de bajo nivel que ofrecemos en términos de complejidad aritmética no es incompatible con algunas de las aportaciones teóricas presentes en la abultada literatura. De hecho, ésta nos permite entender mejor muchas de las discusiones pasadas.

### Factores Contextuales

Hay, además de la complejidad computacional, otros factores a tener en cuenta, al menos con probabilidades simples, aunque también éstos pueden ser explicados usando el mismo principio, los pasos de cálculo necesarios.

Mostrábamos en la primera parte de este trabajo cómo con probabilidades simples la coincidencia o no en escala entre instrucciones y respuesta (*Emparejamiento I-R*) podía determinar los resultados en mayor medida que el formato de representación. La explicación para este fenómeno es relativamente sencilla. En problemas de probabilidad simple, las diferencias que pudieran existir en cuanto a complejidad de cálculo entre probabilidades y frecuencias naturales son mínimas, si no inexistentes. Sin embargo, al no haber un emparejamiento entre instrucciones y respuesta, es necesario realizar un cambio de escala, lo que supone un cálculo y, por lo tanto, un aumento en la complejidad del problema con respecto a aquellas situaciones en las que instrucciones y respuesta coinciden en escala.

Cuadro 18. Pasos de cálculo y Emparejamiento Instrucciones-Respuesta.

**Emparejamiento I-R**

**Instrucciones (I)**

“*la probabilidad de que piquen es del 10%”*

**Respuesta (R)**

*¿Cuál sería la probabilidad de que fuera picante?*  10 %

**I: 10% R: 10%**

**Pasos de cálculo: 0**

**No emparejamiento I-R**

**Instrucciones (I)**

“*la probabilidad de que piquen es de 0,1”*

**Respuesta (R)**

*¿Cuál sería la probabilidad de que fuera picante?*  10 %

**I: 0,1 R: 10%**

**Pasos de cálculo: 1**

(0,1\*100)

Para los proponentes de la *Ecological Rationality*, a lo largo de nuestra historia evolutiva se ha desarrollado un algoritmo cognitivo especializado en el tratamiento de frecuencias naturales. Si existiera dicho módulo, no es descabellado pensar que éstas deberían también mejorar el rendimiento en los problemas de probabilidad simple o, más bien, que el módulo podría directamente favorecer a las frecuencias o probabilidades simples, ya que para aplicar un algoritmo Bayesiano necesitamos haber visto antes varias probabilidades simples. Ya hablamos antes sobre esto. Si el criterio relevante para la aparición de una adaptación cognitiva es Lamarckiano, las probabilidades simples cuentan con una clara ventaja sobre las posteriores.

Nuestros resultados muestran que el formato de representación no parece determinar el rendimiento en problemas de probabilidad simple. El único factor que facilita o dificulta el cálculo es, de nuevo, la complejidad computacional, vestida de emparejamiento entre instrucciones y respuesta.

### Adaptaciones e hipótesis frecuentista

Al proponer algo como una adaptación es importante mostrar que la hipótesis alternativa no puede explicar la evidencia existente (Andrews et al., 2003). Es también importante tener presente la posibilidad de que se trate de una exaptación (un rasgo preexistente adquiere un nuevo efecto beneficioso sin modificación del fenotipo por selección), especialmente en el dominio cognitivo (Andrews et al., 2003), donde, por ejemplo, las redes neuronales presentan una opción en la que un sólo mecanismo de aprendizaje puede subyacer a diferentes sistemas cognitivos (Kruschke, 1992).

El mayor problema de la hipótesis frecuentista, tal y como hemos comentado antes, es la dificultad de desenredarla empíricamente de su alternativa, la hipótesis de la complejidad. Todo el mundo está de acuerdo en que los formatos frecuentistas hacen más sencilla la computación en diversas situaciones (particularmente con los problemas Bayesianos clásicos y enfrentando probabilidades relativas con frecuencias absolutas). La evidencia empírica (relativa a esa comparación) muestra que, en general, los resultados son mejores en problemas Bayesianos expresados en forma de frecuencias naturales. El problema radica en que estos problemas son también más sencillos computacionalmente, así que la evidencia puede ser reformulada a: en general, los resultados son mejores en problemas Bayesianos cuando la complejidad computacional es menor. Así pues, la hipótesis alternativa puede perfectamente explicar la evidencia existente. De hecho, la potencia explicativa de la hipótesis alternativa es mayor, ya que incluye fenómenos como las clases de referencia relativas y absolutas (experimento 2.2.), explica también situaciones en las que el rendimiento es mayor con probabilidades que con frecuencias naturales (ver experimento 2.3.) y, finalmente, predice diferencias entre distintas condiciones de frecuencias naturales (experimento 2.4.) .

Un algoritmo cognitivo del tipo presentado por los frecuentistas debería estar basado en un sistema de representación numérica innato. Sin embargo, hasta ahora, las únicas propuestas de sistemas de representación numérica innatos (Dehaene, 1992; Dehaene, 1997; Dehaene, 2001; Feigenson et al., 2004) son mucho menos sofisticados que los que la hipótesis frecuentista necesitaría. Hay, por supuesto, algunas discrepancias en cuanto a los detalles de las diferentes propuestas teóricas, aunque parece haber un acuerdo generalizado sobre algunas propiedades centrales de la representación numérica. En concreto, que las representaciones numéricas son imprecisas, abstractas y que pueden ser comparadas y combinadas a partir de operaciones aritméticas simples (adición y substracción) (Spelke y Kinzler, 2007).

Dehaene y colaboradores (Dehaene, 1992; Dehaene, 1997; Dehaene, 2001; Feigenson et al., 2004), por ejemplo, proponen dos sistemas centrales de número. El primero, que nos ofrece una representación aproximada de número que captura las interrelaciones entre grupos (*numerosities*)(Feigenson et al., 2004), es el menos polémico. El segundo sirve para mantener un seguimiento preciso de cantidades pequeñas de objetos y para representar información sobre sus propiedades cuantitativas continuas (Feigenson, Carey, y Hauser, 2002; Feigenson et al., 2004), y no es aceptado globalmente. Algunos autores consideran que no existe un módulo numérico específico subyacente a la discriminación entre pequeños grupos de objetos, que el sistema que gestiona pequeñas cantidades lo hace basándose en atributos perceptivos como superficie total (Ni y Zhou, 2005).

En cualquier caso, las conclusiones que trascienden a cualquier polémica sobre cuáles son exactamente los sistemas innatos para trabajar con números que tenemos son que éstos estarían limitados en su poder de representación. No soportarían conceptos como fracciones (probabilidades), raíces cuadradas, números negativos o números enteros exactos (frecuencias naturales). La construcción de números naturales, racionales y reales depende de procesos costosos que probablemente solo son accesibles a humanos educados de un subconjunto de culturas, aunque estos procesos se basan en el sistema o sistemas esenciales que corresponden al sentido numérico básico (Feigenson et al., 2004).

Es difícil para nosotros, con dichas restricciones, pensar en un módulo o sistema innato dedicado a mantener un registro de la frecuencia de los eventos (e.g. 196 de 623), y computar las probabilidades subsiguientes usando frecuencias. Finalmente, el hecho de que esta representación básica de número se haya encontrado en bebés, niños, adultos y primates no humanos (Feigenson et al., 2002; Hauser, 2005; Spelke y Kinzler, 2007; Wynn, Bloom, y Chiang, 2002) hace que la idea de que exista un módulo o algoritmo específico para frecuencias sea especialmente complicada.

Algunos autores han tratado de conjugar los campos de las matemáticas cognitivas (*cognitive mathematics*), representación numérica, etc., con el cálculo de probabilidad Bayesiano (Bonato, Fabbri, Umilta, y Zorzi, 2007; Ni y Zhou, 2005), señalando las dificultades existentes tanto en niños (Mix, Cohen Levine, y Huttenlocher, 1999; Ni y Zhou, 2005) como en adultos (Bonato et al., 2007), a la hora de trabajar con números racionales o fracciones, lo que estaría emparejado con las dificultades a la hora de trabajar con probabilidades.

Parece que todo el mundo está de acuerdo en las dificultades que entraña el uso de números no enteros como fracciones (Bonato et al., 2007; Ni y Zhou, 2005). Está el *sesgo del numero entero* (*Whole-Number Bias*) para atestiguarlo (Ni y Zhou, 2005). Éste consiste en la tendencia de los niños a usar los esquemas básicos de los números enteros para interpretar fracciones. En definitiva, es una prueba más de que los humanos manejan peor fracciones que números enteros. Los porcentajes (probabilidades) son un tipo de fracción (una fracción de 100) y las frecuencias naturales son consideradas como una pareja de números enteros (Bonato et al., 2007; Butterworth, 1999; Ni y Zhou, 2005).

Lo interesante es que, para autores como Butterworth, la superioridad de las frecuencias no tiene nada que ver con la relativa simplicidad computacional de las frecuencias sobre las probabilidades en formulaciones Bayesianas, más bien, tiene que ver con que el cerebro humano ha sido configurado de nacimiento para representar conjuntos y sus magnitudes (Butterworth, 1999, 2001). La explicación que dan, por tanto, para las diferencias entre probabilidades y frecuencias en problemas de cálculo Bayesiano se aleja de la aportada por los propios frecuentistas sobre la menor dificultad del cálculo con frecuencias (Hoffrage, Lindsey, Hertwig, y Gigerenzer, 2001) y redunda en la segunda dimensión frecuentista sobre las raíces evolutivas de la diferencia.

Hay varios problemas en esta otra versión frecuentista. El primero es que, tal y como han mostrado algunos autores, el sistema que representa sets lo hace de manera borrosa e imprecisa (Feigenson et al., 2004). Además de esto, no solo hay evidencia empírica sobre la dimensión de la complejidad aritmética, sino también puramente formal y teórica. Por otro lado, como hemos demostrado en varios de nuestros experimentos, los números enteros (frecuencias) no son siempre más sencillos que los racionales (probabilidades) y, de hecho, su complejidad formal puede ser igualada y comprobada empíricamente.

Por encima de todas las discusiones y discrepancias existentes, la visión de la *Ecological Rationality* ha resultado ser extraordinariamente fructífera. Han presentado una aproximación basada en las “*Simon’s scissors*” para el estudio de la cognición humana (Todd y Gigerenzer, 2002, 2007) en la que las dos hojas son igualmente importantes. Las dos hojas representan la mente humana y el mundo que hay alrededor, destacando al mundo que nos rodea como parte importante e inevitable de la ecuación (Simon, 1990). El problema surge cuando damos por sentado la importancia de una observación ambiental. Es difícil evitar la tentación de usar explicaciones evolutivas plausibles cuando miramos alrededor y vemos frecuencias por todos los lados. Pero, como dijimos antes, es vital mostrar que ninguna alternativa puede explicar la evidencia presente cuando se propone algo como una adaptación (Andrews et al., 2003; Cosmides y Tooby, 1994; Williams, 1996), lo que no es el caso con la hipótesis de las frecuencias naturales como formato favorecido por la evolución (Brase, 2003; Brase y Barbey, 2006; Cosmides y Tooby, 1996; Gigerenzer y Hoffrage, 1995).

El grado de implicación con las tesis evolucionistas de las distintas posibilidades teóricas frecuentistas es variable (Barbey y Sloman, 2007), y aunque, como dijimos antes, esta disección de las posturas evolucionistas es en exceso artificial (Brase, 2007), resulta útil para clarificar las aseveraciones teóricas. Las visiones más radicales (*Mind as Swiss army knife*) basadas en principios modulares estrictos no son aceptadas por el ala evolucionista del grupo frecuentista (Brase, 2007; Ermer, Cosmides, y Tooby, 2007), prefiriendo éstos lo que Barbey y Sloman (2007) llaman heurístico de frecuencias naturales (n*atural frequency heuristic*) (Brase, 2007), que propone que la mente ha evolucionado para procesar frecuencias naturales y que esa adaptación dio lugar al heurístico de frecuencias naturales. Aunque hay ciertas discrepancias entre los frecuentistas a este respecto, ya que mientras Brase indica que el nivel adecuado en el que encajar a los frecuentistas sería el de *natural frequency heuristic* (Brase, 2007), recientemente Gigerenzer rechaza esta opción alineando los heurísticos con aquellas respuestas aproximadas, aunque no propiamente Bayesianas, y minimizando la relación entre los dos ejes en los que se basa la posición frecuentista, la complejidad y el algoritmo Bayesiano innato (Gigerenzer y Hoffrage, 2007). Así que parece que las mismas imprecisiones y malentendidos que han plagado la literatura antifrecuentista, a los que nos referíamos en la introducción, salpican también a las filas frecuentistas.

Las conclusiones que aquí presentamos no dependen realmente del grado concreto de implicación con los principios modularistas. Que nosotros sepamos, no hay ningún defensor de la superioridad de las frecuencias naturales que rechace los argumentos evolutivos. Lo que los resultados empíricos que aportamos muestran es que la hipótesis evolucionista en cualquiera de sus formas parece ser innecesaria. Incompatible en el caso de las versiones más radicales modularistas, y redundante, en cualquier otro. Es incompatible con las versiones más extremas que entienden que existe un módulo impermeable porque, como hemos visto, existen variables cognitivas generales de las que dependen los aciertos a la hora de resolver tareas de cálculo de probabilidad Bayesiano. Y, finalmente, sería innecesaria y redundante en cualquier otro caso porque la hipótesis alternativa basada únicamente en complejidad aritmética entendida como número de pasos de cálculo, no solo es más sencilla y parsimoniosa, sino que nos muestra que la ventaja de las frecuencias naturales se da únicamente en determinados supuestos (Fiedler et al., 2000), puede ser revertida en otros, como hemos visto a lo largo de este trabajo, y, además, no ha sido comprobada en condiciones suficientemente ecológicas (Evans y Elqayam, 2007). Es curioso que con la gran importancia que se le da al muestreo natural en el frente frecuentista, en los experimentos en los que basan sus conclusiones la información se presenta de un modo que nunca podría ser calificado como tal (Evans y Elqayam, 2007).

### ¿Es la mente un sistema general de cómputo o un conjunto de algoritmos o módulos especializados?

El debate sobre formatos de representación de la información trasciende al cálculo de probabilidad y se adentra en parajes mucho más densos, en las profundidades de cómo funciona la mente, o mejor, cómo está estructurada. En la introducción hablábamos sobre la *Ecological Rationality* y cómo desde ella se proponía una visión de la mente regida por un principio modular muy claro. Enfrentada a ésta situábamos a los Heurísticos y Sesgos como ejemplo de teoría basada en un principio radicalmente opuesto, la mente como una máquina de procesamiento de información de propósito general.

La *Ecological Rationality* ve, por lo tanto, a la mente como un conjunto de piezas diseñadas por selección natural para tratar con distintos aspectos de la realidad. La historia filogenética nos ha moldeado a imagen y semejanza del entorno, acoplando funciones adaptativas (a través de algoritmos o módulos especializados) que ofrecen respuestas extraordinariamente ajustadas para aquellas situaciones que de otro modo serían prácticamente intratables. Uno de los algoritmos cognitivos que se han incorporado a nuestro arsenal hace que tratar problemas de cálculo Bayesiano resulte trivial cuando la información está presente en forma de frecuencias naturales. Otras formas de representación como las probabilidades o frecuencias relativas resultan inadecuadas para este algoritmo especializado y los resultados derivados de su uso son inadecuados (Gigerenzer y Hoffrage, 1995).

Nuestros resultados no aportan pruebas sobre cómo está estructurada la mente en general, aunque realmente ése no ha sido en ningún momento el objetivo de esta tesis. Sí nos hablan, sin embargo, sobre el debate particular que motivó este trabajo: ¿Existe un modulo o algoritmo cognitivo especializado para el tratamiento de las frecuencias naturales? Nuestra respuesta es que no. Todas las predicciones y resultados derivadas de la hipótesis frecuentista pueden ser explicadas a partir del principio más parsimonioso de complejidad computacional o aritmética. Y no solo eso, sino que hay, como hemos visto, correlaciones importantes entre pruebas de carácter cognitivo y resultados en problemas clásicos de cálculo Bayesiano con frecuencias naturales, además de varios errores y omisiones en aspectos centrales de la teoría frecuentista (ignorar la importancia de las clases de referencia, predicciones contradictorias con respecto a los menús de la información...). Con un panorama así, resulta completamente innecesario y hasta inoportuno introducir valoraciones evolucionistas que, por interesantes que sean, resultan completamente irrelevantes de cara a discernir la manera en la que los seres humanos se enfrentan al cálculo de probabilidad Bayesiano.

### A modo de conclusión

En cualquier caso, y a modo de conclusión, la principal hipótesis frecuentista no se sustenta bajo el peso de nuestros experimentos. Mostramos, en primer lugar y con probabilidades simples, que factores contextuales (*Emparejamiento I-R*) pueden tener una influencia significativa mientras que el formato de representación no. En segundo lugar, que la complejidad aritmética es un mejor predictor que el formato de representación y, en tercer y último lugar, que existen correlaciones entre medidas cognitivas y el rendimiento de los sujetos en las tareas clásicas de cálculo de probabilidad Bayesiano. La maleabilidad de los resultados al alterar sutiles factores cosméticos, o al variar cualquiera de los componentes de los problemas, y la posibilidad derivada de estas modificaciones de controlar de manera extremadamente precisa los resultados a partir de estas manipulaciones, así como que hipótesis alternativas nos permitan explicar con mayor precisión la realidad, debería leerse como una advertencia en contra de la hipótesis frecuentista (evolutiva) fuerte y moderar las aseveraciones sobre la superioridad intrínseca de las frecuencias naturales como formato de representación.

Hemos pretendido demostrar o, en el peor de los casos, insinuar, que la discusión sobre la superioridad de las frecuencias naturales como formato de representación de la información, que ha generado un debate tan complejo como difícil de seguir en los últimos años, es estéril. Creemos que el extraordinario nivel de contradicciones internas y malentendidos no es casual. Gigerenzer y Hoffrage (1995) propusieron que dicha superioridad se sustentaba sobre dos ejes, complejidad y evolución. Una lectura en profundidad por parte de los revisores del que aquí consideramos artículo fundacional del movimiento frecuentista (Gigerenzer y Hoffrage, 1995) hubiera ahorrado cantidades ingentes de tinta y tiempo. Pero no queremos que se nos malinterprete. Algunas de las aportaciones de dicho artículo son realmente brillantes, aunque hay varias inconsistencias y, sobre todo, afirmaciones extraordinariamente aventuradas, especialmente en lo referente al segundo eje, el evolutivo. Éste es el que ha centrado el mayor volumen de críticas, curiosamente a partir de manipulaciones del primer eje, el de la complejidad.

Y es aquí donde han tropezado gran parte de los críticos. La complejidad conceptual de las ideas planteadas por Gigerenzer y Hoffrage es tal, que un leve despiste convertía un experimento brillante en una mera replicación. Las distinciones entre probabilidades de eventos simples, frecuencias relativas, absolutas, naturales, los significados de manipular las muestras totales, las probabilidades previas y un largo etcétera, han sido usadas como armas arrojadizas para defenderse de errores de bulto, imperdonables en algunos casos, y para crear muros invisibles basados en meras conjeturas.

Por poner algunos ejemplos, son imperdonables los errores que han cometido varios autores al confundir cualquier tipo de frecuencia con frecuencias naturales, al comparar frecuencias naturales con normalizadas (Hoffrage et al., 2002; ver Johnson-Laird et al., 1999; Macchi y Mosconi, 1998), cuando uno de los rasgos básicos de las frecuencias naturales es la no normalización de éstas. Por otro lado, la defensa que realizan los frecuentistas ante los brillantes resultados de Fiedler y colegas (Fiedler et al., 2000) al diferenciar entre escala de referencia igual y distinta, se basa en un hecho del que ni tan siquiera existía comprobación empírica, la influencia de la muestra total al facilitar la conversión de probabilidades a frecuencias (Hoffrage et al., 2002). Hemos intentado comprobar ese extremo en el experimento 2.2. y el resultado ha sido que sigue sin haber ninguna prueba empírica de la influencia de este factor y, por lo tanto, las críticas a Fiedler no parecen tener base.

Podríamos separar en varias dimensiones las aportaciones de este trabajo. En primer lugar, hemos visto cómo la diferencia entre probabilidades y frecuencias naturales se da por las diferencias en complejidad de ambos formatos. Al usar probabilidades simples estas diferencias desaparecen. Cualquier factor que añada complejidad a la resolución de un problema (sea éste con probabilidades simples o posteriores) afecta al rendimiento de los sujetos independientemente del formato de representación. Comprobamos este extremo con probabilidades simples a través de la correspondencia en escala entre instrucciones y respuesta. En probabilidades posteriores con la clase de referencia de la información (absoluta o relativa).

En segundo lugar, la influencia teórica que la presencia o ausencia de la muestra total tenía sobre los resultados, y que sirvió, como decíamos, para criticar algunos trabajos, así como la presencia de la clase complementaria que decían podría ser usada para computar la muestra total y “naturalizar” probabilidades, no tuvo ningún peso significativo en nuestros resultados, lo que debería obligar a matizar algunas afirmaciones o, como mínimo, basarse en comprobaciones empíricas para desechar o criticar resultados.

En tercer lugar, también encontramos que las habilidades cognitivas generales parecen influir en la correcta resolución de los problemas Bayesianos, siendo esta influencia mayor en frecuencias naturales que en probabilidades relativas, lo que, desde luego, no apoya la existencia de un módulo o algoritmo especializado para las frecuencias naturales.

Para acabar, hemos podido comprobar que la complejidad computacional o aritmética, entendida como pasos de cálculo necesarios para resolver el problema, es una herramienta de gran utilidad para predecir el comportamiento de los sujetos y explicar diferencias basadas en formatos de representación, clases de referencia y estructura del problema, entre otros.

# Bibliografía

Andrews, P. W., Gangestad, S. W., y Matthews, D. (2003). Adaptationism–how to carry out an exaptationist program. *Behavioral and Brain Sciences, 25*(04), 489-504.

Bar-Hillel, M. (1977). The base-rate fallacy in probability judgments.

Bar-Hillel, M., y Fischhoff, B. (1981). When do base rates affect predictions?

Barbey, A., y Sloman, S. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences, 30*(03), 241-254.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica, 22*(1), 23-36.

Bishop, C. (1997). *Neural networks for pattern recognition*: Oxford University Press.

Bonato, M., Fabbri, S., Umilta, C., y Zorzi, M. (2007). The mental representation of numerical fractions: Real or integer? *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1410.

Brase, G. (2003). There is no evidentiary silver bullet for the frequency adaptation hypothesis. *Behavioral and Brain Sciences, 25*(04), 508-509.

Brase, G. (2007). Omissions, conflations, and false dichotomies: Conceptual and empirical problems with the barbey and sloman account. *Behavioral and Brain Sciences, 30*(03), 258-259.

Brase, G., y Barbey, A. (2006). Mental representations of statistical information. *Advances in psychology research, 41*, 91–113.

Brighton, H. (2006). *Robust inference with simple cognitive models.* Paper presented at the Between a rock and a hard place: Cognitive science principles meet AI-hard problems, Menlo Park, CA.

Butterworth, B. (1999). *The mathematical brain*. London: Macmillan.

Butterworth, B. (2001). Statistics: What seems natural?, *Science* (Vol. 292, pp. 853-855).

Camerer, C. (1990). Do markets correct biases in probability judgment? Evidence from market experiments. *Advances in Behavioral Economics, Volume 2*, 126.

Chase, V., Hertwig, R., y Gigerenzer, G. (1998). Visions of rationality. *Trends in Cognitive Sciences, 2*(6), 206-214.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated. *Behavioral and Brain Sciences, 4*(3), 317-329.

Cosmides, L., y Tooby, J. (1994). Beyond intuition and instinct blindness - toward an evolutionarily rigorous cognitive science. *Cognition, 50*(1-3), 41-77.

Cosmides, L., y Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58*(1), 1-73.

Daston, L. J. (1980). Probabilistic expectation and rationality in classical probability theory in papers in honor of erwin n. Hiebert. *Historia Mathematica Toronto, 7*(3), 234-260.

De Finetti, B. (1989). Probabilism. *Erkenntnis, 31*(2), 169-223.

De Neys, W. (2007). Nested sets and base-rate neglect: Two types of reasoning? *Behavioral and Brain Sciences, 30*(03), 260-261.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*(1-2), 1-42.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*: Oxford University Press.

Dehaene, S. (2001). Precis of the number sense. *Mind & Language, 16*(1), 16-36.

Eddy, D. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. *Judgment under uncertainty: Heuristics and biases*, 249-267.

Einhorn, H. J., y Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgement and choice. *Annual Reviews in Psychology, 32*(1), 53-88.

Ermer, E., Cosmides, L., y Tooby, J. (2007). Functional specialization and the adaptationist program. *The Evolution of Mind: Fundamental Questions and Controversies*.

Estes, W. K., Campbell, J. A., Hatsopoulos, N., y Hurwitz, J. B. (1989). Base-rate effects in category learning: A comparison of parallel network and memory storage-retrieval models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(4), 556-571.

Evans, J., y Elqayam, S. (2007). Dual-processing explains base-rate neglect, but which dual-process theory and how? *Behavioral and Brain Sciences, 30*(03), 261-262.

Evans, J., Handley, S., Perham, N., Over, D., y Thompson, V. (2000). Frequency versus probability formats in statistical word problems. *Cognition, 77*(3), 197-213.

Feigenson, L., Carey, S., y Hauser, M. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological Science, 13*(2), 150-156.

Feigenson, L., Dehaene, S., y Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307-314.

Fiedler, K., Brinkmann, B., Betsch, T., y Wild, B. (2000). A sampling approach to biases in conditional probability judgments: Beyond base rate neglect and statistical format. *Journal of Experimental Psychology General, 129*(3), 399-418.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25-42.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond “heuristics and biases”. *European Review of Social Psychology, 2*(1), 83-115.

Gigerenzer, G. (1993). The bounded rationality of probabilistic mental models. *Rationality: Psychological and philosophical perspectives*, 284-313.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). *Subjective Probability*, 129–161.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to kahneman and tversky. *Psychological Review, 103*(3), 592-596.

Gigerenzer, G. (2002). *Reckoning with risk: Learning to live with uncertainty*. London: Penguin Books.

Gigerenzer, G., y Brighton, H. (2007). Can hunches be rational. *Journal of Law, Economics and Policy, 4*(1), 155-176.

Gigerenzer, G., y Fiedler, K. (2003). Minds in environments: The potential of an ecological approach to cognition. *Manuscript submitted for publication.[GG]*.

Gigerenzer, G., Hell, W., y Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 513-525.

Gigerenzer, G., y Hoffrage, U. (1995). How to improve bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*(4), 684-704.

Gigerenzer, G., y Hoffrage, U. (2007). The role of representation in bayesian reasoning: Correcting common misconceptions. *Behavioral and Brain Sciences, 30*(03), 264-267.

Gigerenzer, G., Hoffrage, U., y Kleinbölting, H. (1991). Probabilistic mental models: A brunswikian theory of confidence. *Psychological Review, 98*(4), 506-528.

Gigerenzer, G., y Murray, D. (1987). Cognition as intuitive statistics: Hillsdale, NJ: Erlbam.

Gigerenzer, G., y Selten, R. (2001). *Bounded rationality: The adaptive toolbox*: MIT Press.

Girotto, V., y Gonzalez, M. (2001). Solving probabilistic and statistical problems: A matter of information structure and question form. *Cognition, 78*(3), 247-276.

Girotto, V., y Gonzalez, M. (2002). Chances and frequencies in probabilistic reasoning: Rejoinder to hoffrage, gigerenzer, krauss, and martignon. *Cognition, 84*(3), 353-359.

Gluck, M. A., y Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117*(3), 227-247.

Gould, S., y Lewontin, R. (1979). The spandrels of san marco and the panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B, Biological Sciences (1934-1990), 205*(1161), 581-598.

Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *Quart. J. Econ, 95*, 537–557.

Grether, D. M. (1992). Testing bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior and Organization, 17*(1), 31–57.

Griffin, D., y Dukeshire, S. (1993). The role of visual random sampling in base rate use and neglect: A methodological critique. *Unpublished manuscript, University of Waterloo, Waterloo, Ontario, Canada*.

Hacking, I. (1975). *The emergence of probability : A philosophical study of early ideas about probability, induction and statistical inference*. London; NewYork: Cambridge University Press.

Hamm, R. M. (1993). Explanations for common responses to the blue/green cab probabilistic inference word problem. *Psychological reports, 72*(1), 219-242.

Hauser, M. (2005). Our chimpanzee mind. *Nature, 7055*, 60.

Hertwig, R., y Gigerenzer, G. (1999). Theconjunction fallacy revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making, 12*, 275-305.

Hinton, G. E., McClelland, J. L., y Rumelhart, D. E. (1986). Distributed representations. In *Mit press computational models of cognition and perception series* (pp. 77-109).

Hoffrage, U., Gigerenzer, G., Krauss, S., y Martignon, L. (2002). Representation facilitates reasoning: What natural frequencies are and what they are not. *Cognition, 84*(3), 343-352.

Hoffrage, U., Lindsey, S., Hertwig, R., y Gigerenzer, G. (2001). Response: Statistics: What seems natural?, *Science* (Vol. 292, pp. 853-855).

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Legrenzi, M. S., y Caverni, J. P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review, 106*(1), 62-88.

Kahneman, D., Slovic, P., y Tversky, A. (1982). Evidential impact on base rates. In A. Tversky y D. Kahneman (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 153-160).

Kahneman, D., y Tversky, A. (1979). Prospect theory - analysis of decision under risk. *Econometrica, 47*(2), 263-291.

Kahneman, D., y Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103*, 582-591.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences, 19*(1), 1-53.

Krauss, S., Martignon, L., Hoffrage, U., y Gigerenzer, G. (2002). Bayesian reasoning and natural frequencies: A generalization to complex situations. *Manuscript submitted for publication*.

Kruschke, J. K. (1992). Alcove: An exemplar-based connectionist model of category learning. *Psychological Review, 99*(1), 22–44.

Landman, J., y Manis, M. (1983). Social cognition: Some historical and theoretical perspectives. *Advances in experimental social psychology, 16*, 49-123.

Laplace, P. S. (1995). *Philosophical essay on probabilities* (F. W. Truscott y F. L. Emory, Trans.): Springer.

Lewontin, R. (1978). Adaptation. *Scientific American, 239*(3), 213-230.

Lewontin, R. (1979). Sociobiology as an adaptationist program. *Behav Sci, 24*(1), 5-14.

Lopes, L. L. (1991). The rhetoric of irrationality. *Theory & Psychology, 1*(1), 65.

Macchi, L. (2000). Partitive formulation of information in probabilistic problems: Beyond heuristics and frequency format explanations. *Organizational Behavior and Human Decision Processes, 82*(2), 217-236.

Macchi, L., y Mosconi, G. (1998). Computational features vs frequentist phrasing in the base-rate fallacy. *Swiss Journal of Psychology, 57*, 79-85.

Macdonald, R. R. (1986). Credible conceptions and implausible probabilities. *British Journal of Mathematical & Statistical Psychology, 39*, 15-27.

Mix, K. S., Cohen Levine, S., y Huttenlocher, J. (1999). Early fraction calculation ability. *Developmental Psychology, 35*, 164-174.

Ni, Y., y Zhou, Y. D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist, 40*(1), 27-52.

Nosofsky, R. M., Kruschke, J. K., y McKinley, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 211–233.

O’Brien, D., Roazzi, A., y Maria da Graça, B. B. (2004). Raciocínio sobre probabilidades condicionais: As evidências a favor da hipótese freqüentista se fundamentam em comparações errôneas. *Estudos de Psicologia, 9*(1), 35-43.

Roazzi, A., O'Brien, D. P., y Dias, M. (2003). On the frequentist and probabilistic debate:" Dumb-luck" Turns out to be a plausible explanation. *Psicologia: Reflexão e Crítica, 16*, 389-402.

Simon, H. A. (1990). Invariants of human behavior. *Annual Reviews in Psychology, 41*(1), 1-20.

Sloman, S., Over, D., Slovak, L., y Stibel, J. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes, 91*(2), 296-309.

Spelke, E. S., y Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89-96.

Todd, P., y Gigerenzer, G. (2000). Precis of simple heuristics that make us smart', behavioral and brain sciences, 23 (5): October.

Todd, P., y Gigerenzer, G. (2002). Shepard's mirrors or simon's scissors? *Behavioral and Brain Sciences, 24*(04), 704-705.

Todd, P., y Gigerenzer, G. (2003). Bounding rationality to the world. *Journal of Economic Psychology, 24*(2), 143-165.

Todd, P., y Gigerenzer, G. (2007). Environments that make us smart: Ecological rationality. *Current Directions in Psychological Science, 16*(3), 167-171.

Tversky, A., y Kahneman, D. (1974). Judgment under uncertainty - heuristics and biases. *Science, 185*(4157), 1124-1131.

Tversky, A., y Kahneman, D. (1982). Evidential impact of base rates in judgment under uncertainty, eds. Daniel kahneman, paul slovic, & amos tversky: Cambridge: Cambridge University Press.

Tversky, A., y Kahneman, D. (1983). Extensional versus intuitive reasoning - the conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293-315.

Villejoubert, G., y Mandel, D. (2002). The inverse fallacy: An account of deviations from bayess theorem and the additivity principle. *Memory & Cognition, 30*(2), 171-178.

Williams, G. (1996). *Adaptation and natural selection: A critique of some current evolutionary thought*: Princeton University Press.

Wynn, K., Bloom, P., y Chiang, W. C. (2002). Enumeration of collective entities by 5-month-old infants. *Cognition, 83*(3), 55-62.

Zhu, L., y Gigerenzer, G. (2006). Children can solve bayesian problems: The role of representation in mental computation. *Cognition, 98*(3), 287-308.