Comparing Measures of Individual Differences in Performance of Conditional Reasoning*

Gloria Seoane, María Dolores Valiña, María José Ferraces & Montserrat Martín

University of Santiago de Compostela (Spain)

Introduction

The study of thinking and reasoning is a topic of central interest for economists, anthropologists, logicians, pedagogues and of course for psychologists. A central problem in the experimental investigation in Psychology is to describe how people think and reason deductively and inductively.

There are three fundamental theoretical approaches to deductive reasoning in the Cognitive Psychology: mental logic, mental models and pragmatic schemas.

---

* This work was presented at 10th *EUROPEAN MEETING OF THE PSYCHOMETRIC SOCIETY*, celebrated in Santiago de Compostela, Spain (15-18th July, 1997).

There are several proponents of a universal mental logic (Inhelder & Piaget, 1958) or natural logics (Braine, 1978, 1990, 1994; Braine & O´Brien, 1991; Braine & Rumain, 1981, 1983; Osherson, 1974, 1975; Rips, 1983, 1990, 1994). Other authors propose that reasoning is based on construction and evaluation of mental models (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991). A third approach asserts that reasoning is not based on general inference rules and assumes that people have domain-specific reasoning mechanisms such as pragmatic reasoning schemas inductively acquired (Cheng & Holyoak, 1985, 1989; Cheng, Holyoak, Nisbett & Oliver, 1986; Holyoak & Cheng, 1995) or innates procedures for identify potential deviations from social contracts (Cosmides 1985, 1989; Cosmides & Tooby, 1992).

Psychometrics studies the thinking from a different perspective. The central interest for the researchers in Psychometrics is not the understanding of underlying cognitive processes and mental representations but the study of the individual differences in these mental processes.

However, despite the differences between these two approaches to the study of human reasoning the categorical syllogisms and linear syllogisms were included on early intelligence tests (Burt, 1919, 1921; Guilford, 1959; Thurstone, 1938). Moreover, in the past decades there is a novel and comparatively neglected field: the study of qualitative and quantitative differences in reasoning. Roberts (1993, p. 575) suggested that:

"*The problem of individual differences is as follows: if a theory of reasoning is being proposed that is intended to describe the processes used by all people for all reasoning tasks, then what is the status of this theory if it is subsequently found that not all people are using the same processes?*."

Galotti, Baron & Sabini (1986) examined the correlates of reasoning ability on a syllogistic reasoning task. They found evidence for the use of both models and rules of reasoning. In a previous work Sternberg and Weil (1980) found individual differences in reasoning strategies (a mental model strategy, a deduction rule strategy and a mixture of both) in the resolution of experimental tasks that involve linear syllogisms.

Alternatively, Sternberg and Gastel (1989) investigated information processing during the solution of inductive reasoning problems (analogies, classifications and series completions) and also administered five psychometric tests to each subject. They showed correlations between experimental tasks and psychometric tests. These correlations address two principal questions:

First, are scores on the experimental tasks related to scores on the psychometric tests?. Second, do the correlations with the reasoning tests differ from those with verbal/perceptual factor?. It was found that the correlations of the experimental task with the reasoning tasks are higher than those with verbal/perceptual tasks. Thus, *"the experimental tasks do appear to tap abilities related to those tapped by the psychometric tests"* (p. 8).

Despite the importance of conditional reasoning in daily life, the study of qualitative individual differences has not become a central focus in cognitive or psychometric studies. There is no nearly previous experimental research about this issue. We found in a previous study (Valiña, Seoane, Ferraces & Martín, 1995) a considerably better performance in the Wason´s selection task in the higher verbal group (DAT-VR) but there were no differences between subjects with high and low scores on the PMA-E psychometric test. In the present experiment we explore the relation among different measures

in psychometric ability tests (verbal comprehension and reasoning) and the performance of this experimental conditional reasoning task.

The experimental task.

The Wason´s selection task is one paradigm widely used for studying conditional reasoning. The original problem was elaborated by Wason (1966, 1968). He presented a conditional rule: "*very card that has a vowel on one side has an even number on the other*" and four cards: E, K, 4 and 7. The subjects´ task is to decide which cards should be turned over to test the conditional rule.

Frequently, the subjects only selected the E card (*p*) or the E and 4 cards (*p* and *q*). The correct response is the selection of the E and 7 cards (*p* and *not q*), but only 5-10% of the subjects chose these cards. The subjects selected a case for which the rule is true, but it is a negative instance which provides a violating case and can prove the truth or the falsity of the rule.

We selected this task in part because has long been of interest to experimental psychologist (see Evans, 1982, 1984, 1989; Evans, Newstead & Byrne, 1993; Newstead & Evans, 1995, Wason, 1983, for revisions) and moreover because it is one of the most widely used paradigm for studying the importance of factors related to the role of pragmatic knowledge in reasoning (Chrostowski & Griggs, 1985; Girotto, Gilly, Blaye & Light, 1989; Griggs, 1983, 1989; Johnson-Laird, Legrenzi & Legrenzi, 1972; Manktelow & Over, 1991; Pollard & Evans, 1987; Valiña, Seoane, Ferraces & Martín, 1995, 1996; Wason, 1983; Wason & Shapiro, 1971; Yachanin, 1986).

Concretely, in this paper we examined the following questions: (1) the relation among different measures in psychometric ability tests (verbal

comprehension and reasoning), computerized measure of comprehension skills and the subjects´ performance in the experimental task of conditional reasoning, (2) whether or not good and poor comprehenders sistematically differ in their performance in Wason´s selection task (Wason, 1966, 1968) and (3) the differential influence of rule content and instructions on the subjects´ performance in this selection task.

METHOD

Subjects

One hundred and fifty-four undergraduates (20 males, 134 females; mean age 21 years), studying Psychology at the University of Santiago de Compostela (Spain) collaborated in this study. The students participated as partial fulfillment of a course requirement. They didn´t perform a similar experiment before and none had any prior training in formal logic.

Data from 18 participants were not used because they failed to follow the experimental instructions or they had not completed all the task.

Materials and apparatus

1) Psychometric tests

The participants completed three spanish versions of three psychometric ability tests: DAT-VR, PMA-V and PMA-R.

USC

2) Gernsbacher´s Comprehension Battery

The spanish version of the Battery was presented on a DX-486 computer using a computer programme elaborated by Manuel de Vega, of the University of La Laguna (Spain). The programme presents 4 narrative texts with times of exposition on the screen of 3,5 seconds for each sentence. Subjects read sentences that were presented sentence-by-sentence on the computer monitor. After the last sentence of each story disappeared, a test of five alternative questions about each experimental text appeared. The time of presentation of each question about the story were of 20 seconds. The subjects´ task was to select as rapidly and accurately as possible the correct alternative that had occurred in the text they had just finished reading. Finally, the programme presented the next text 15 seconds after the final response of the subject.

The programme registered both the correct responses and the reaction times of the participants to the questions about the stories.

3) Selection task

Each subject received three rules, with the following types of content: abstract, thematic-permission and thematic-obligation. Half of the subjects received true-false instructions of the rule and the other half violation instructions. The test booklets were used in previous investigations (Martín, 1996; Valiña et al., 1996). The information for each of the three tasks was as follows:

a) *Abstract selection task*. "If a Wasit card has an A on one side, then it must have a 3 on the other". The four cards presented to the subjects were: "A", "K", "3" and "7".

b) *Thematic-permission*. In this rule a law was expressed; therefore it is similar to permission. The rule was: "If a person is more than 18 years old, then he has the right to vote". The four cards said: "20 years old", "16 years old", "you have the right to vote" and "you do not have the right to vote".

c) *Thematic-obligation*. The rule expressed a traffic regulation: "If a person rides a motorbike, then they must wear a helmet". The four cards that were represented were: "motorbike", "car", "helmet" and "cap".

The instructions were used previously (Chrotowski & Griggs, 1985; Valiña et al., 1995, 1996; Yachanin, 1986). In the *true-false version*, the instructions were:

"Your task consists of selecting cards and only those that must be turned over to decide if the rule is true or false (select those cards which you consider necessary to turn over to check if the person carrying out the experiment has lied or not in relation to the composition of the rule".

For the *violation version*, the instructions were:

"Your task consists of selecting only those cards that must be turned over in order to decide if the rule is being violated or not".

Two different versions were made for each of the types of booklets. In one of these the thematic versions were at the beginning, followed by the abstract rule and in the other the abstract version was included at the beginning. Additionally, the order of presentation of the two thematic versions was counterbalanced.

Procedure

Participants met in groups of up to 12 with two experimenters over 2 days. On the 1st day they received in the laboratory both the psychometric tests with conventional instructions and the spanish version of the Gernsbacher´s Comprehension Battery. Subjects were tested with each interacting on a separate microcomputer in the same laboratory.

On the 2nd day of the experiment, participants were assigned at random to one of two experimental groups: (1) true-false instructions and (2) violation instructions. Subjects were tested in groups of 12. Each subject received a booklet with instructions on the first page, followed by three selection tasks (an abstract one and two thematic selection tasks). The instructions were read to the subjects and questions were solicited to ensure that they understood the instructions. Finally, they were instructed to work at their own rhytm, without a time limit.

RESULTS

The analysis were carried out with the data from the 136 subjects, once those who had not completed the task had been eliminated.

First we performed ANOVAs to test the differential influence of rule content and instruction on the subjects´ performance in the selection task (Wason, 1966, 1968). Other ANOVAs were performed for checking whether or not good and poor comprehenders sistematically differ in their execution in Wason´s selection task, and finally we performed analysis in order to provide a test of the relation among different measures in psychometric

ability tests (verbal comprehension and reasoning), the computerized measure of comprehension skills and the subjects´ performance in the experimental task of conditional inference.

1) ANOVAS

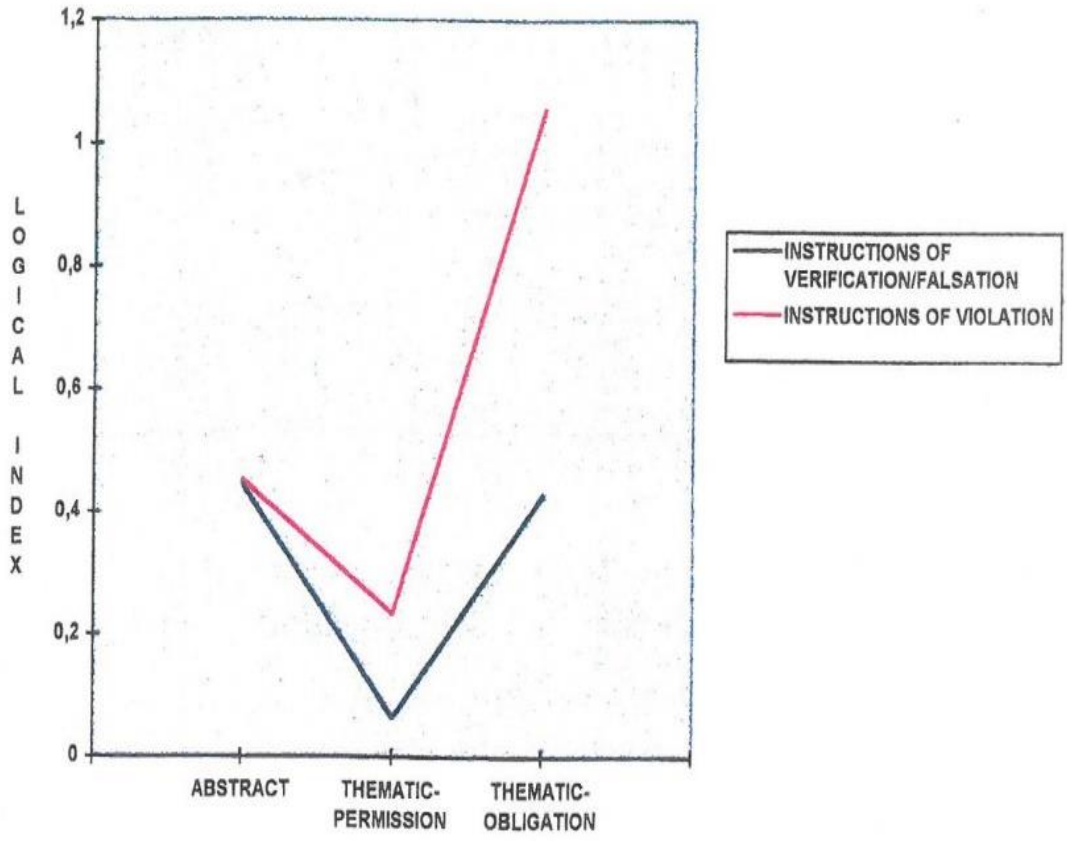
The logical and matching indices were calculated for each of the three tasks. Both indices vary between +2 and -2, according to Pollard and Evans (1987). In the logical index the *p* or *not-q* selection gave a mark of +1 and the *not-p* or *q* selection gave -1. In the matching index the *p* or *q* selection gave a mark of +1 and with -1 the *not-p* or *not-q* selection. ANOVAS 2 x 3 (instructions x content) were made for each type of index, with the data from the 136 participants.

1.a) Logical index

For the logical index the principal effects of the *content* ($F_{(1.82;\ 244.52)}$= 21.61; $p <$ .0001; ε = .912) and the *instructions* ($F_{(1,\ 134)}$ = 6.59; p <.011) were registered. In the thematic-obligation higher logical indices were obtained (M=.765), followed by the abstract version of the task (M=.449) and the thematic-permission (M =.154). Similarly, the logical indices were higher in those subjects who received violation instructions (M= .87) compared with those who received true-false instructions (M=.31).

Significant interactive effects have also been registered for *instructions x content* ($F_{(1.82;\ 244.52)}$ = 6.32; $p <$ .003; ε =.912). In the thematic-obligation task higher indices were obtained in those subjects who received violation instructions (see figure 1).

FIGURE 1. INTERACTIVE EFFECTS BETWEEN INSTRUCTIONS AND CONTENT IN THE LOGICAL INDEX



1.b) Matching index

A significant effect was obtained with the *instructions* ($F_{(1,\ 134)} = 11.31$; $p < .001$). Participants obtained higher matching index with instructions for checking the rule (M=1.042) compared to those who received violation instructions (M=.58).

Similarly, significant effects were registered in the *content* ($F_{(2,\ 268)}$ = 25.64; p <.0001). Concretely, the highest matching index was obtained with abstract content (M =1.11), followed by the thematic-obligation (M =.897) and the thematic-permission (M =.375). Abstract content differs significantly from the other groups ($F_{(1,134)}$ = 24.42; p < .0001) and similarly thematic-permission differs significantly from the thematic-obligation ($F_{(1,\ 134)}$ = 26.97; p < .0001) by orthogonal tests.

2) Three-way mixed ANOVAS

Because the central questions being addressed involve group differences, we performed three-way mixed analyses of variance, with group (good vs. poor verbal comprehenders / good vs. poor reasoners), instructions (true-false vs. violation instructions) and content (abstract, thematic-permission and thematic-obligation) as factors, with repeated measures on the last factor.

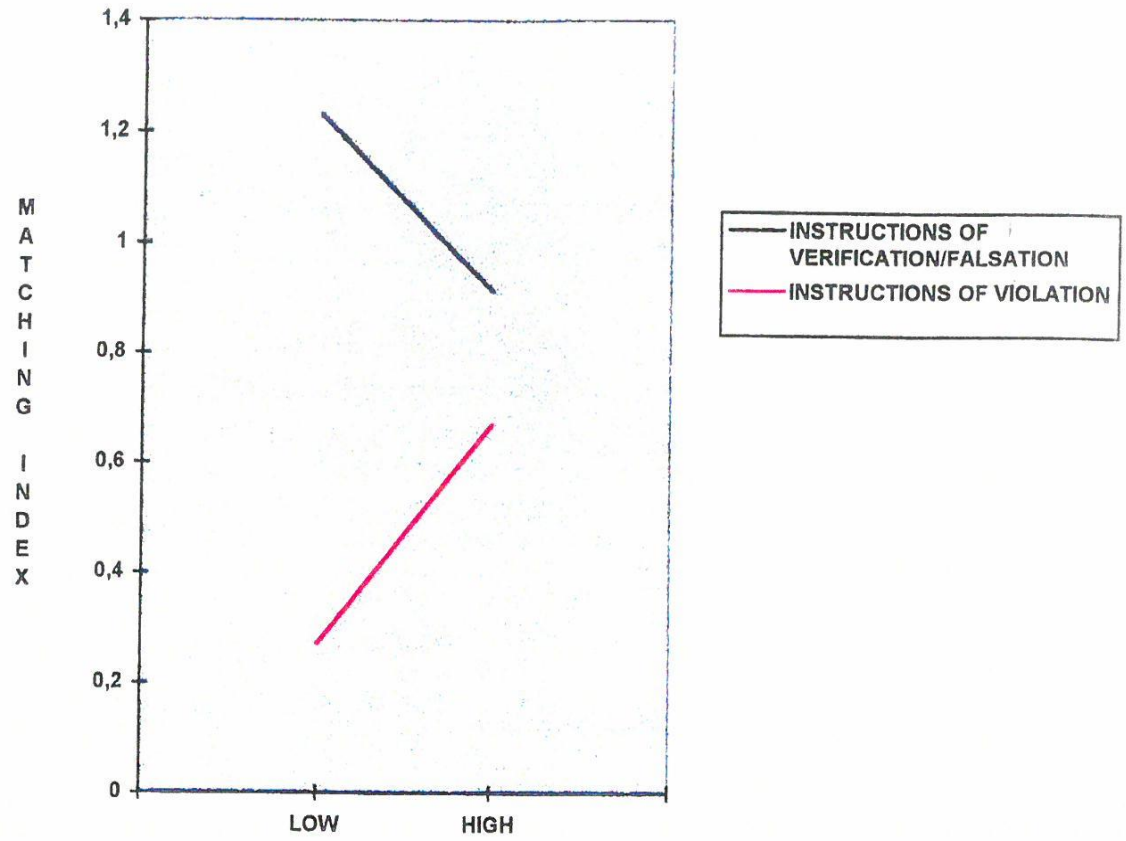
In terms of differential analyses there were no differences in the logical and matching indices among good and poor verbal comprehenders (PMA-V & Gernsbacher´s Comprehension Battery) or subjects with high and low scores in the PMA-R, but there were significant differences among good and poor reasoners (DAT-VR).

The logical index was considerably better (M=1.131) in the higher reasoning-verbal group vs. the group with low scores in the DAT-VR (M = .386). The differences were significant ($F_{(1,\ 70)}$ = 8.52; p < .005).

Similarly, for the matching index the interaction between *group and instructions* was significant ($F_{(1,70)}$=5.02; p <.028. There were no differences in the matching index of good reasoners in function of the experimental instructions, but differences in the group with low scores in the DAT-VR

were found. Particularly, the poor reasoners obtained highest matching indices with instructions for checking the rule (see figure 2).

FIGURE 2. INTERACTIVE EFFECTS BETWEEN GROUP AND INSTRUCTIONS IN THE MATCHING INDEX.



3) CORRELATIONS OF THE EXPERIMENTAL TASK WITH THE PSYCHOMETRIC TESTS SCORES

We performed another analysis to provide a test of the relation among different measures in verbal and comprehension psychometric tests, the computerized measures of comprehension skills and the subjects´ performance in the experimental task with both logical and matching indices.

The analysis were carried out (a) with the data from the total of 136 subjects and (b) with de data from the two experimental groups (true-false instructions and violation instructions).

The results show for the total sample (N=136) that: (a) the scores of the DAT-VR are related with the performance in Wason´s selection task with the abstract content (r = .317; *p* = .0001) and with the thematic-permission (r = .2656 ; *p* = .002) in terms of logical index and (b) there is a significant relation between scores in the DAT-VR psychometric test and the computerized measures of Gernsbacher´s Comprehension Battery (r = .1663; *p* = .05).

With the *true-false instructions* (N=63) there was found a significant relation between scores in the DAT-VR and the logical index (r = .3524; *p* = .005) and the matching index (r = - .3779; *p* = .002) with the abstract content. There was also a significant relation between measures in the Gernsbacher´s Comprehension Battery and the logical index for this abstract content (r = .2862; *p* = .023).

With *the violation instructions* (N = 73) there was a significant relation between measures in the DAT-VR and the logical index for the abstract content (r = .2939; *p* = .012) and the thematic-permission content (r = .4316; *p* = .0001).

CONCLUSIONS

From this study we draw the following conclusions:

1.- A good predictor for performance in Wason´s selection task was the DAT-VR psychometric test.

2.- There were differences in performance among subjects with high and low scores in the DAT-VR. The logical index was considerably better in the higher reasoning verbal-group.

3.- The poor reasoners in the DAT-VR obtained highest matching indices with instructions for checking the rule. There were no differences for the subjects with high scores in this psychometric task.

4.- There were no differences in the selection task performance between groups with scores high and low in the PMA-V, PMA-R psychometric tests and the Gernsbacher´s Comprehension Battery.

5.- The highest logical index was registered with thematic-obligation content, followed by the abstract content and the thematic-permission content.

6.- The highest matching index was obtained with abstract content, followed by the thematic-obligation and finally the lowest matching index was obtained with the thematic-permission content.

7.- The logical indices are higher in those subjects who received violation instructions with the thematic-obligation content.

Acknowledgments

We especially thanks to Manuel de Vega, professor of the University of La Laguna (Tenerife, Spain), for providing us his Spanish computerized version of Gernsbacher´s Comprehension Battery.

References

Braine, M.D.S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, *85*, 1-21.

Braine, M.D.S. (1990). The "natural logic" approach to reasoning. In W.F. Overton (Ed.). *Reasoning, neccesity and logic: Developmental perspectives* (pp. 135-157). Hillsdale, NJ: Erlbaum.

Braine, M.D.S. (1994). Mental logic and how to discover. In J. Macnamara, & G.E. Reyes (Eds.). *The logical foundations of cognition* (pp. 241-263). Oxford, England: Oxford University Press.

Braine, M.D.S., & O´Brien, D.P. (1991). A theory of If: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, *98*, 182-203.

Braine, M.D.S., & Rumain, B. (1981). Development of comprehension of "or". Evidence for a sequence of competencies. *Journal of Experimental Child Psychology*, *31*, 46-78.

Braine, M.D.S., & Rumain, B. (1983). Logical reasoning. In J.H. Flavell & E.M. Markman (Eds.). *Handbook of child Psychology: 3. Cognitive Development* (pp. 263-340). New York: Wiley.

Burt, C. (1919). The development of reasoning in school children. *Journal of Experimental Pedagogy*, *5*, 68-77.

Burt, C. (1921). *Mental and scholastic tests*. London: King & Son.

Cheng, P.W., & Holyoak, K.J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391-416.

Cheng, P.W., & Holyoak, K.J. (1989). On the natural selection of reasoning theories. *Cognition, 33,* 285-313.

Cheng, P.W., Holyoak, K.J., Nisbett, R.E., & Oliver, L.M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18,* 293-328.

Chrostowski, J.J., & Griggs, R.A. (1985). The effects of problem content, instructions and verbalization procedures on Wason´s selection task. *Current Psychological Research and Reviews, 4,* 99-107.

Cosmides, L. (1985). *Deduction or darwinian algorithms? An explanation of the "elusive" content effect on the Wason selection task*. Doctoral Dissertation, Harvard University. University Microfilms 86-02206.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31,* 187-276.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J.H. Barkow, L. Cosmides, & J. Tooby (Eds.). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 163-228). New York: Oxford.

Evans, J.St.B.T. (1982). *The psychology of deductive reasoning*. London: Routledge and Kegan Paul.

Evans, J.St.B.T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology, 75,* 451-468.

Evans, J.St.B.T. (1989). *Bias in human reasoning: Causes and consequences*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Evans, J.St.B.T., Newstead, S.E., & Byrne, R.M.J. (1993). *Human reasoning: The Psychology of deduction*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Galotti, K.M., Baron, J., & Sabini, J. (1986). Individual differences in syllogistic reasoning: Deduction rules or mental models? *Journal of Experimental Psychology: General*, *115*, 16-25.

Girotto, V., Gilly, M., Blaye, A., & Light, P. (1989). Children´s performance in the selection task: Plausibility and familiarity. *British Journal of Psychology*, *80*, 79-85.

Griggs, R.A. (1983). The role of problem content in the selection task and THOG problem. In J. St.B.T. Evans (Ed.). *Thinking and Reasoning: Psychological Approaches* (pp. 16-43). London: Routledge and Kegan Paul.

Griggs, R.A. (1989). To "see" or not to "see": That is the selection task. *The Quarterly Journal of Experimental Psychology*, *41A*, 517-529.

Guilford, J.P. (1959). Three faces of intellect. *American Psychologist*, *14*, 469-479.

Holyoak, K.J., & Cheng, P.W. (1995). Pragmatic reasoning about human voluntary action: Evidence from Wason´s selection task. In S.E. Newstead, & J.St.B.T. Evans (Eds.). *Perspectives on thinking and reasoning. Essays in honour of Peter Wason* (pp. 67-89). Hove: Lawrence Erlbaum Associates, Inc.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. London: Routledge and Kegan Paul.

Johnson-Laird, P.N. (1983). *Mental models. Towards a cognitive science of language, inference and consciousness*. Cambridge, England: Cambridge University Press.

Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction*. Hove, U.K.: Lawrence Erlbaum Associates.

Johnson-Laird, P.N., Legrenzi, P., & Legrenzi, S. (1972). Reasoning and a sense of reality. *British Journal of Psychology*, *63*, 336-400.

Manktelow, K.I., & Over, D.E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition*, *39*, 85-105.

Martín, M. (1996). *Una exploración del razonamiento cotidiano: Importancia del conocimiento en inferencia condicional*. Unpublished doctoral dissertation. University of Santiago de Compostela, Spain.

Newstead, S.E., & Evans, J.St.B.T. (1995). *Perspectives on thinking and reasoning. Essays in honour of Peter Wason*. LEA. Hillsdale.

Osherson, D.N. (1974). *Logical abilities in children. Vol.2. Logical inference: Underlying operations*. Potomac, MD: Lawrence Erlbaum Associates Inc.

Osherson, D.N. (1975). *Logical abilities in children. Vol.3. Reasoning in adolescence: Deductive inference*. Hillsdadle, NJ: Lawrence Erlbaum Associates Inc.

Pollard, P., & Evans, J.St.B.T. (1987). On the relationship between content and context effects in reasoning. *American Journal of Psychology*, *100*, 41-60.

Rips, L.J. (1983). Cognitive Processes in propositional reasoning. *Psychological Review*, *90*, 38-71.

Rips, L.J. (1990). Reasoning. *Annual Review of Psychology*, *41*, 321-353.

Rips, L.J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: MIT Press.

Roberts, M.J. (1993). Human Reasoning: Deduction Rules or Mental Models, or Both?. *The Quarterly Journal of Experimental Psychology, 46A*(4), 569-589.

Sternberg, R.J., & Gastel, J. (1989). If dancers ate their shoes: Inductive reasoning, with factual and counterfactual premises. *Memory & Cognition*, *17* (1), 1-10.

Sternberg, R.J., & Weil, E.M. (1980). An aptitude x strategy interaction in linear syllogistic reasoning. *Journal of Educational Psychology*, *72*(2), 226-239.

Thurstone, L.L. (1938). Primary mental abilities. *Psychometric Monographs*, *1*. (Adpt. in Tea, 1972).

Valiña, M.D., Seoane, G., Ferraces, M.J., & Martín, M. (1995). Tarea de selección de Wason: Un estudio de las diferencias individuales. *Psicothema*, *7*(3), 641-653.

Valiña, M.D., Seoane, G., Ferraces, M.J., & Martín, M. (1996). Wason´s selection task: Content effect, instruction effect or both?. In *Abstracts of The Third International Conference on Thinking* (p. 11). British Psychological Society. Cognitive Psychology Section. University College London.

Wason, P.C. (1966). Reasoning. In B.M. Foss (Ed.). *New horizons in Psychology, I* (pp. 135-151). Harmondsworth: Penguin.

Wason, P.C. (1968). Reasoning about a rule. *The Quarterly Journal of Experimental Psychology*, *20*, 273-281.

Wason, P.C. (1983). Realism and rationality in the selection task. In J.St.B.T. Evans (Ed.). *Thinking and reasoning: Psychological Approaches* (pp. 44-75). London: Routledge and Kegan Paul.

Wason, P.C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *The Quarterly Journal of Experimental Psychology*, *23*, 63-71.

Yachanin, S.A. (1986). Facilitation in Wason´s selection task: Content and instructions. *Current Psychological Research and Reviews*, *5*, 20-29.