Current Psychology
https://doi.org/10.1007/s12144-021-01966-3

Content of relationship, number of alternatives and working memory capacity in conditional inferences

Cristian A. Rojas-Barahona[1] · Sergio Moreno-Ríos[2] · Juan A. García-Madruga[3] · Carla E. Förster[4]

Accepted: 3 June 2021

Abstract
This study evaluates the contribution of conditional relationship contents to making inferences, depending on the span of the participant's working memory. This is achieved through the use of a deductive reasoning task with conditional statements and the "Reading Span Test" by Daneman and Carpenter, in its Spanish version. The results of the two experiments show that the execution of an inference task is mediated by participants' working memory span. WM is correlated with the most complex valid inferences, but not with the most automatic ones. Contents and the number of alternatives to the antecedent only had an effect on the inference of the high-WM group. The results are relevant in several ways, and can help us to refine our predictions about inferences.

Keywords Conditional relationship contents · Attribute relationships · Working memory

Nowadays, there is a consensus about the key role of working memory (WM), both in the explanation of how people solve complex cognitive tasks (e.g., see Bjorklund, 2005; Engle et al., 1999) and in the intellectual changes that occur over the course of development (e.g., see Alloway & Copello, 2013; Gathercole et al., 2004). Thus, the main theories about deductive reasoning, such as heuristic probabilistic approaches (Chater & Oaksford, 1999; Copeland & Radvansky, 2007), formal inference rules (Braine, 1978; Braine & O'Brein, 1991, 1998; Rips, 1994) and mental models (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991, 2002), all share the notion that both intelligence (e.g., see Stanovich & West, 2008) and WM capacity (e.g., see De Neys et al., 2005; Markovits et al., 2002), influence the deductive process.

WM is conceived as a limited-capacity system that allows individuals to store temporary information, thus enabling them to perform complex cognitive tasks, such as learning, understanding, and reasoning (Baddeley, 1986). One of the most widely studied WM models, proposed by Baddeley and Hitch (1974), posits the existence of a system known as the "central executive": a limited-capacity attentional system, in charge of controlling and coordinating subordinate systems, the "phonological loop" and the "visuo-spatial sketchpad". The former is responsible for codifying and storing language-based information, whereas the latter stores and manipulates visuo-spatial images. A fourth element of the model is the "episodic buffer", thought to be an interface between the central executive and long-term memory (Baddeley, 2000), and this could have an important role in the influence of emotion on WM (Baddeley et al., 2009). The central executive has been considered to be a key element accounting for the differences observed in the execution of reasoning tasks (see Klauer et al., 1997). Among the primary functions of the central executive, besides the capacity to focus, divide and switch attention, a fourth main function is the activation of representations in LTM (see, for instance, Baddeley, 2007).

The Relationship between Deductive Reasoning and Working Memory

Theories of mental rules propose the existence of a set of formal rules (for example, "if P then Q" or "P ⊃ Q". "P"; therefore "Q": the Modus Ponens rule), which make it possible to infer a conclusion through premises. These mental rules

✉ Cristian A. Rojas-Barahona
c.rojas@utalca.cl

[1] Faculty of Psychology, Universidad de Talca, Av. Lircay S/N, Campus Lircay, Talca, Chile

[2] Escuela de Psicología, Universidad de Granada, Granada, Spain

[3] Escuela de Psicología, UNED, Madrid, Spain

[4] Vicerrectoría de Pregrado, Universidad de Talca, Talca, Chile

Published online: 17 June 2021

are believed to act at a syntactic level, separate from content and context ("P" and "Q" could have any content, and "Q" would always be derived from "P"). The deductive process is thought to involve three phases: 1) interpretation of the statement, which leads to the logical form of the argument (for example, the statement "If it is cold, then it is raining. It is cold" could take the logical form "P ⊃ Q. P"); 2) access to an appropriate repertoire of mental rules required to arrive at the conclusion of the problem (in this case, the Modus Ponens rule, which leads to "Q"); and 3) translation of the formal conclusion into the contents of the problem (in this case, "Therefore, it is raining"). From this perspective, it has been pointed out that some inferences are easier than others. Studies have proposed the existence of direct and automatic inferences alongside indirect or strategic ones (Braine & O'Brein, 1991; Braine et al., 1984). It has been pointed out that effort and WM load are directly proportional to the steps required by a deductive action: more steps involve more effort and a heavier WM load. Two logical inferences that display a clear difference in the number of steps required are Modus Ponens (MP), in which the antecedent of the statement is affirmed, and Modus Tollens (MT), in which the consequent is denied. The former is regarded as very easy. In this case, the MP rule would be stored in the mind; any problem containing this rule would simply activate it (intuitive). The case of MT is thought to require the use of several rules to obtain the correct conclusion, the MP rule and "*reductio ad absurdum*" (not fully intuitive, see Ricco et al., 2020). Although Evans and Over (2004) do not accept the vision of an innate mental logic, their suppositional theory also assumes that a MT inference is performed via the suppositional strategy of *reductio ad absurdum*, not required by a MP inference (see Barrouillet et al., 2008).

The mental model theory posits that, when reasoning, individuals construct models in the form of semantic representations analogous to the situations of the world represented (see Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991, 2002). That is, given the conditional "If there is a pen, then there is a notebook", individuals construct a mental representation (maybe an image) which includes a pen and a notebook. There are three deductive steps: first, understanding the statement by creating a representation. For example, understanding this statement leads to the construction of the following representation:

pen notebook
....

where the first line represents an explicit model or possibility which includes both the 'pen' and the 'notebook', while the second (….) represents an implicit model or mental note which informs the individual that understanding the statement implies other possibilities or mental models. From this point of view, in order to avoid overloading their working memory, reasoners begin by constructing the smallest number of models possible. In the second stage, based on the information provided by the categorical premise (There is a pen) and by the model constructed, an informative conclusion is formulated by matching the new information with that included in the model; in the previous example, the conclusion is:

Therefore: there is a notebook

Finally, in order to test this conclusion, reasoners search for alternative models by constructing new ones. Thus, participants who reason more thoroughly may also construct the other two possibilities or models implied by the statement "If there is a pen, then there is a notebook", and thus arrive at a complete representation such as the following:

pen notebook
non-pen notebook
non-pen non-notebook

Due to the limited capacity of WM, the mental model theory predicts that the more mental models that are required to reach a conclusion, the higher the difficulty of a given reasoning task. Considering alternative mental models takes longer than just one. In this regard, and as in the previous approach, there are thought to be differences between MP and MT inferences. MP can be solved without alternative models. Conversely, for MT, other models are necessary to construct a full representation like the one above, which results in a higher WM load and longer time for making an inference.

However, inferences are not always made analytically. In contexts involving temporal pressure (Evans & Curtis-Holmes, 2005) or high WM loads (De Neys, 2006), participants tend to give more automatic answers, based on accepting conclusions consistent with their beliefs. Chater and Oaksford (1999) postulate that participants approach problems from the probability theory rather than by using logic.

Two Types of Processing

Two different ways of processing have been proposed: one intuitive and fast, called "processing Type 1", and the other deliberate and slow, called "processing Type 2" (e.g. see Evans & Stanovich, 2013). WM has become crucial in the distinction between the two types of processing: Type 1 processing requires minimal demands on WM while Type 2 processing is strongly based on what Stanovich (2011) calls "cognitive decoupling", based on the WM operation that is a key feature for System 2. The distinction between the two systems or types of processing (questioned by some authors, such as

Kruglanski, 2013) can be integrated in each of the main theories of deduction. For all of them, the MP inference can be made automatically, and is, therefore, akin to the Type 1 processing proposal, while MT inferences require in addition a controlled Type 2 processing. For example, Khemlania and Johnson-Laird (2012) proposed that the idea of Type 1 and Type 2 processing corresponds to the inferences that can be made with the initial model, such as the MP inference, and those that require looking for alternatives (the fleshing-out operation), such as the MT inference. Accordingly, MP inferences are expected to be faster and less error-prone than MT inferences.

It is clear that the theories presented suggest a connection between WM capacity and the difficulty of a given reasoning task, and predict that MT inference will be more difficult than MP inference. Thus, one would expect a positive correlation between WM capacity and the ability to make MT inferences but not MP inferences, given the low difficulty of the latter. Consistent with this hypothesis, Toms et al. (1993) observed that the frequency of acceptance of MT inferences in a conditional reasoning task was affected in association with WM load. However, Markovits et al. (2002) obtained different results. This study considered valid MP and MT inferences, as well as the fallacies of affirming the consequent (AC) and denying the antecedent (DA), which, respectively, result from accepting the affirmation of the antecedent or the denial of the consequent as conclusions. The researchers observed a positive correlation between the MP, AC, and DA inferences and WM capacity, but did not find evidence to support a similar phenomenon in the case of MT inferences. Therefore, the empirical data gathered is not as clear as prior theoretical assumptions.

The Importance of Content in Deductive Inference

Individuals are willing, in general, to accept inferences depending on the content of the premises (Vandierendonck et al., 2006). Theories intended to account for this phenomenon assume the existence of different explanatory mechanisms. The mental model theory holds that certain contents make it easier to evaluate possibilities than do others (Johnson-Laird & Byrne, 2002). According to the mental rule theories and the suppositional theory, some contents produce pragmatic effects which lead to the application of "guest" inferences, not present in the logical arguments (Braine & O'Brein, 1991; Evans & Over, 2004; Rips, 1994). Content may modify the conditional probability that the consequent emerges when the antecedent is presented (Evans et al., 2003). Yet, what happens if we use arbitrary contents with statements such as "if there is a sheet of paper, it is red" and compare them with "if there is a sheet of paper, there is a pen"? In this case, the inference mechanism should work equally, regardless of the nature of the proposed representation.

Thus, an inference based on a conditional should be made in the same way when the relationship between the antecedent and the consequent refers to an attribute of a single object, "a sheet of paper is red", or to relationships between two different objects, "a sheet of paper and a pen". In contrast, Markovits (2000) proposed that, based on the mental model theory, the semantic structure of the individual, generated by his/her experience of the world, influences the representation of the contents of the conditional and conditions the inferences made. Markovits and Barrouillet (2002) suggested that during the construction of the conditional representation, a relationship is activated between the antecedent and the consequent depending on the degree of association present in the structure of the individual's knowledge. Consistent with the proposals advanced by Anderson and Lebiere (1998), Cantor and Engle (1993) and Cowan (2001), the authors stated that the elements with the most activation force will be retrieved from memory and incorporated into the individual's representation. Therefore, participants would be expected to perform differently in conditional reasoning tasks depending on whether relationships of attribute or between-objects are expressed.

With respect to WM, Cowan (2001) stressed the importance of activation force in memory retrieval, and the relationship between the limits of the focus of attention and the restriction of the number of WM "chunks". This could be one of the reasons behind the differences observed in deductive tasks. Considering the limits of WM and the differences between individuals' WM capacity when making more "cognitive effort" (Sperber & Wilson, 1986) – due to the need to represent and/or manipulate more chunks – participants would be expected to display differences in their deductive actions. If we regard chunks as a set of concepts strongly associated with each other, it could be said that, depending on the type of relationship between the contents of a conditional (strong or weak), it should be easier to associate or dissociate a chunk, which will result in more or less cognitive effort.

Thus, if the proposal by Markovits and Barrouillet is correct, there should be differences between making inferences when the connection of the antecedent with the consequent of the conditional refers to an attribute (stronger relationship) and when it refers to two different objects (weaker relationship). For example, in the case of an MT inference: "If there is a triangle, it is red", "it is not red", the initial representation would be discarded (the red triangle or the supposition of the triangle) and an alternative would be sought. The operation refers to a single object. Similarly, when there is a relationship between-objects, "If there is a triangle, there is a square", "there is not a square", the initial representation ("the triangle and the square") would be discarded, and alternatives would be sought. In this case, the statement refers to two objects. The relationship between the triangle and its red attribute is stronger than that between the triangle and the square. The action of connecting different objects, as in the case of MP inferences, would be expected to result in a higher WM load than when the relationship refers to an attribute. Similarly, higher degrees of connection between the antecedent and the consequent in complex inferences requiring the dissociation of elements, such as in MT inferences, would also increase WM demand.

If the previous is correct, we would expect different ways of processing the valid inferences between participants with high WM capacity and those with low WM capacity. Participants with higher WM capacity are able to consider alternatives, which is time consuming. Therefore after "if there is a triangle it is red; it is not red …" participants with a higher WM capacity might think about "squares, circles…" while participants with lower WM would just conclude "not triangle". In a typical inference task, the correct conclusion to select is "not a triangle". Therefore, we would predict that participants with higher and lower WM will give the correct conclusion with the same frequency, but the former would take longer. In this study we will consider both measures as indices of processing: time spent in processing and the frequency of responses. Easier inferences will be faster and/or more correct.

In brief, this study evaluates two novel predictions applied to participants with high working memory capacity, but not those with low working memory: deduction is easier for attribute than for between-objects conditionals and easier for single than for multiple alternatives conditionals. Also based on previous findings, with the new conditional contents, we expect to find that: MP is easier than MT, and therefore faster and more accurate responses are expected. Complex inferences (MT), but not simple ones (MP) are easier for people with higher WM. We will concretise these predictions below.

Experiment 1

In this study, we evaluated the possible influence of semantic structure as expressed in a conditional relationship related to a property of the same entity or to a relationship between different properties in participants with high working memory. In order to do this, we compared the execution of simple MP inferences with more complex MT ones in an inference task. Some authors have highlighted the necessity for deductive theories to incorporate semantic structure into long-term memory to predict differences when making logical inferences (Barrouillet, 2011; Barrouillet, & Lecas, 1999, 2002; Markovits, 2014).

An MT inference requires the reasoner to discard his/her initial representation. In an MT inference, the consequent is denied and the correct logical answer consists in denying the antecedent. This type of inference can be made just by discarding the antecedent (no p) or by looking for an alternative concrete antecedent. Participants with higher WM capacity are expected to be better able to compute alternatives than participants with low WM capacity. Also, the relation between the antecedent and the consequent could make the operation more difficult. Consider the following two examples:

> If there is a triangle, then there is a square. There is not a square. Therefore, there is not a triangle.
> If there is a triangle, then it is red. It is not red. Therefore, there is not a triangle.

Although the structure is identical, the contents are different. There is no apparent reason to propose that one inference is easier than the other. There seem to be no differential pragmatic factors involved, and consequently, no differences are expected from mental rule theories. The conditioned probability of the consequent, given the antecedent, does not seem to justify any a priori differences. Therefore, no differences are expected either from the suppositional theory, either, nor from other probabilistic approaches. We could expect a similar prediction of mental models, but the theory establishes that objects, properties and relationships are part of the mental model, and the two cases mentioned are not equivalent. There is something that makes the two conditionals different: one involves two entities (triangle and square) while the second involves only one entity (triangle) and an associated property (red). The negation of the property in MT might make it easier to discard the entity if that semantic structure is part of the deductive process.

If the MP inference is automatic, the characteristics of semantic information will have little influence on it. On the other hand, it is likely to influence alternative-seeking processes, so it should only have an effect on MT inference, not on MP inference. WM capacity may facilitate or limit the differential processing of complex MT inferences. Therefore, we expect that logical MT inferences, but not MP inferences, will display a positive correlation with the WM test. The latter prediction emerges from the notion that, in order to respond correctly to a MT inference, it is necessary to flesh out the mental models, or use more mental rules, while an MP inference only requires an initial model or a mental rule.

Some people might be making the complex inferences based on heuristic processes that do not rely on the WM (e.g., see Evans, 1993; Klaczynski, 2001). One possibility is that people make the MT not in an analytic way but by applying some heuristic process. For example, they could just conclude, not q, when not p is presented, regardless of the content of the premises. If this is the case, no differences produced by the content or by the working memory capacity of participants would be expected.

To summarise, following predictions derived from the deductive theories we predict that:

Table 1 Predictions of the deductive theories about MP and MT inferences for High and Low working memory groups and the effect on relationship (Attribute & Between-objects)

WM Groups	Effect predicted	Mental Rule theories	Suppositional theory	Mental model theory
All	Difficulty (MT>MP)	Yes	Yes	Yes
Low	Relationship effect on MP	No	No	No
High	Relationship effect on MP	No	No	Yes Attribute harder than between-objects

(Prediction-1) MP inferences are easier than MT
(Prediction-2) a correlation between MT and WM capacity but not with MP and
(Prediction-3) Based on the mental model theory, we predict an effect of contents only in MT inferences and only in participants with the highest WM capacity: more difficult inferences for MT attribute contents than between-objects contents (see Table 1).

Method

Participants

A total of 61 university students, 9 male and 52 female, were evaluated. Their average age was 20 years and 8 months (age range: from 17 years and 11 months to 37 years and 4 months). All participants belonged to the Spanish-speaking University of Granada and agreed to participate voluntarily through informed consent. The project was reviewed and approved by the ethics committee of the University of Granada.

Design and Materials

The independent variables considered were: logical inferences (MP and MT), content of the relationship (attribute and between-objects), and WM group (low or high) and the dependent variables were accuracy (correct response) and latency (includes time of the statement, the premise and the answer) of the participants' answers. To evaluate WM capacity, the classic Reading Span Test (Daneman & Carpenter, 1980) was used in its Spanish version (Elosúa et al., 1996). Also, a deductive reasoning task with conditionals was constructed in a computer version (E-Prime) that recorded the accuracy and latency of the participants' answers. The "inferences" and "content of the relationship" variables were manipulated within-participants, while an inter-group approach was used for the "WM group" variable. To separate groups, the extreme groups design presented by Conway et al. (2005) was used, dividing the participants into quartiles and selecting the lower and upper quartiles according to their score on the RST, after ensuring it had a normal distribution. The scoring criterion was <2.8 for the low WM group (19 participants) and > 3.4 for the high WM group (18 participants).

Reading Span Test (RST)

The version used was created by Elosúa et al. (1996), based on the RST constructed by Daneman and Carpenter (1980). This is a double-task test: reading phrases (processing) and remembering the last word of the phrases read (memory). Since the tasks require some kind of attentional control besides a WM span measure, RST is also a measure of the central executive in working memory (see Engle & Oransky, 1999; Whitney et al., 2001). The software displays text on a white background. In the task, participants must remember the last word in a series of phrases previously read aloud. The series starts with two phrases and can reach a maximum of five. Each level presents three sets or series of phrases: three two-phrase series, three three-phrase series, three four-phrase series, etc. For the participant to become familiar with the task, immediately after the instructions, he/she is shown a few sample phrases. There is no time limit to answer.

The following are examples of the phrases the participants read:

> "The noise coming from the street was so loud that we had to go to another place".
> "Our mother baked a chocolate cake for us to take to the party"

Immediately afterwards, the participant had to remember and say aloud, in the order that they appeared, the last words of each phrase: "place" and "party", and do so for each of the series.

To score the RST, the strict criteria applied in the original work of Daneman and Carpenter (1980) were used (e.g., see García-Madruga et al., 2007).

Conditional Reasoning Task

A computer version (E-Prime v2.0 software; Psychology Software Tools, 2007), which measures the accuracy and latency of the participants' answers, was used. The task includes instructions with a practice example, and 20 arbitrary contents of factual conditional statements with their respective categorical premises and options for answers. Sixteen valid logical inferences (MP and MT) and four invalid logical inferences (AC and NA) are included (see Appendix 1). Fallacies were included so answers were not all affirmative or negative. The contents of the conditional statements were distributed homogeneously (attribute or between-objects). Questions were asked at random. Two different versions were presented: the same content was applied with MP in the first version and with MT in the second. Also, the order of the positive and negative alternatives was distributed randomly.

The instructions were the following:

"This is a test to measure how we reason, not an intelligence test. You will have to answer 20 questions, choosing ONE of the 3 alternatives presented. To pick the correct alternative, you must press the number to its left. You must consider that all the information provided is always true. It is important that you answer as quickly as possible. The following is an example:

If it is a knife, then it is made of wood
Imagine it is a knife
Therefore,
<1 > it is not made of wood.
<2 > it is made of wood.
<3 > there is no possible conclusion.

In this case, the correct answer was: IT IS MADE OF WOOD, therefore, you had to press 2. Now, you will solve a question which is very similar to the one you will complete in the experimental phase.

<press SPACE to start the trial>

(The previous example is repeated at this point. The difference is that the participant must press the key associated with the correct alternative)

The experimental phase will begin now. Remember that you must answer as quickly as possible.

<to start, press SPACE>"

Procedure

All the participants completed the two tasks in individual sessions, which lasted approximately 25 min, in a private room at their university. The tasks were presented in the following order: 1) Conditional Reasoning Task and 2) RST. The two tasks were conducted using a computer. The participant sat at the computer, while the researcher sat to his/her left and both started reading the instructions together. After reading all the instructions, the researcher asked if there were any questions. When the instructions were clear, the tests started. In the Conditional Reasoning Task, the software (E- Prime) recorded the accuracy and latency of the participants' answers. Since the participants could start inferring once the categorical premise was presented, latencies were measured considering the total reading time of the categorical premise plus the time taken to choose a conclusion. The inferences were presented in random order. In this test, each participant was shown one of the two versions constructed. In the RST, the researcher recorded all the answers verbalised by the participant, using a specially designed form.

Results and Discussion

Two variance analyses (repeated measures ANOVA) were conducted, one with accuracy values and the other with latencies, considering the following factors: Type of inference (MP or MT) x Content (Attribute or Between-objects) x WM Group (High or Low). Data were analysed using SPSS 22.0 software. Table 1 shows the answers provided by the participants.

Analysis of Accuracy

The variance analysis showed that only the "Type of inference" factor was significant, thus confirming (Prediction 1) the difference in difficulty of MP (m = 97%) and MT (m = 76%) ($F(1, 35) = 26.471$, $p = 0.000$, partial $\eta2 = 0.431$) with a higher frequency for MP (see Table 2).

Analysis of Latency

A logarithmic transform was performed to analyse the latency of correct answers. The variance analysis (repeated measures ANOVA) revealed differences in the "Type of inference" factor ($F(1, 28) = 43.685$, $p = 0.000$, partial $\eta2 = 0.609$); the response times for MT (m = 8.54) were higher than for MP (m = 8.33) confirming Prediction 1. In addition, the second grade interaction Inferences * Content * WM Group ($F(1, 28) = 4.118$, $p = 0.050$, partial $\eta2 = 0.128$) in which the response times were observed, was significant, both with low WM ($F(1, 13) = 32.77$, $p = 0.00$, partial $\eta2 = 0.716$) and high WM groups ($F(1, 15) = 15.000$, $p = 0.001$, partial $\eta2 = 0.500$), who spent more time on MT than MP; in addition, confirming Prediction 3, in the high WM group, in MT only, the time for inferences with attribute contents took longer (8.64) than with between-objects contents (8.54) ($F(1, 16) = 12.418$, $p = 0.003$, partial $\eta2 = 0.437$) (see Table 3). Also the interaction Content * WM group ($F(1, 28) = 4.277$, $p = 0.050$, partial

Table 2 Percentage of MP and MT answers, in general and in the two WM groups (low and high), according to the contents of the relationship (Attribute and Between-objects), Experiment 1: Correct answers are in bold

General (*N*=61)	Modus Ponens Attribute	Modus Tollens Attribute	Modus Ponens Between-objects	Modus Tollens Between-objects
Yes	**96**	14	**98**	17
No	3	**77**	2	**74**
It is unknown	1	9	0	9
Low WM (*N*=19)				
Yes	**95**	26	**98**	23
No	3	**69**	2	**66**
It is unknown	2	5	0	11
High WM (*N*=18)				
Yes	**97**	9	**97**	13
No	3	**78**	3	**74**
It is unknown	0	13	0	13

$\eta2 = 0.126$), in which the high WM group had more latency than the low WM group, was significant on "attribute" ($F(1, 28) = 43.685$, $p = 0.000$, partial $\eta2 = 0.609$) but on "between object" there were no differences ($F(1, 28) = 0.003$, $p = 0.957$, partial $\eta2 = 0.000$).

Correlations

Confirming Prediction 2, no correlations were observed between the percentage of correct responses for logical inferences in MP and the RST ($r = 0.07$, $p = 0.583$). However, there was a correlation between accuracy measures in MT and the WM test ($r = 0.24$, $p = 0.012$).

The higher frequency of MP inferences compared with MT ones and their lower response times are consistent with the results of previous studies and with the theories of reasoning that suggest a more automatic processing of the former (Prediction 1): based on an initial model, an innate rule, or resulting from simulation. It does not require any additional operations: no extra models, rules or steps need be applied. This result can be observed in both the high and low WM groups. Were it associated with WM, individuals with higher capacities should be able to make this inference more easily (Prediction 2). The results of the correlation analysis point in this direction: there was no significant correlation between correct MP responses and RST. The correlation between MT and WM capacity (according to the RST), however, replicates the results presented by Toms et al. (1993), while it counters those by Markovits et al. (2002).

Regarding Prediction 3, manipulating the type of relationship between the antecedent and the consequent should not affect the probability of the presence of one or the other, nor should the number of models constructed or the inference rules applied; however, it is relevant for the construction of representations. If the inference (*reductio ad absurdum* or the use of models) is made through representations with content, the invitation to integrate information (in the case of attribute information) should have some degree of influence on MT inferences. If we are told "if there is a triangle, it is red" and then "it is not red", dismantling the object that we are thinking about (discarding the triangle) should be more difficult than in the inference "if there is a triangle, there is a square, there is not a square." If the information is integrated, more time must be spent dismantling the object while inferring. The results obtained do not clearly show that this is so. They do, however, show that a group of participants (those with more WM capacity) treat conditionals differently.

It may be that participants who displayed high WM capacities, when dealing with MT inferences and attribute-related contents, create an integrated representation (red triangle)

Table 3 Mean (Standard Deviation) of latency in seconds of correct answers to MP and MT inferences in low and high WM groups, according to the contents of the relationship (Attribute and Between-objects), Experiment 1

	MP (attribute)	MT (attribute)	MP (betw. obj.)	MT (betw. obj.)
General	8.30(.25)	8.56(.35)	8.35(.29)	8.51(.30)
Low WM	8.23(.25)	8.47(.37)	8.24(.23)	8.48(.36)
High WM	8.36(.24)	8.64(.32)	8.45(.27)	8.54(.24)

which they must dismantle when they consider the categorical premise (it is not red). This should take longer, but it does not benefit them at all, due to the characteristics of the task. WM span may result from the use of strategies which make information reduction more effective (better coding, leaving more space for information processing; for example, see Case, 1981, 1985). If they used a different strategy, it did not help.

What would have happened with this manipulation of the relationship between the elements expressed in the conditional if the task had incorporated a manipulation with alternatives, which could be multiple? In this case, it would have been necessary to search for counterexamples to complete the task successfully, which would have favoured individuals with higher WM capacities. Single-alternative consequents should be easier to process than multiple-alternative ones (see Schroyens et al., 2000).

Experiment 2

Study 1 showed that high WM individuals did not perform better than low WM participants in MT inferences. The proposed explanation is that when evaluating conclusions ("if p then q" "no q", therefore…), there are only two alternatives with explicit denials (no p or p). The task is too simple, even for people who use strategies that do not involve evaluating all the available alternatives. Therefore, correct responses could include those from participants who used a heuristic strategy and those using a controlled strategy. The new task allows us to distinguish between those responses.

One of the reasons why Barrouillet and Lecas (2002) support including representational structure in long-term memory is the fact that when we make inferences with conditionals which include binary contents (with one alternative; for example "If the lights are *on*, the door is *open*"), we can easily consider the denied alternatives and block the processing of other alternatives. One of the most noteworthy results is obtained when using implicit negative premises (blue) instead of explicit ones (is not red), conditional inferences decreasing (Evans & Handley, 1999; Schaeken, & Schroyens, 2000), except when using natural binary items ("the lights are on", instead of "the lights are not off") (Schroyens et al., 2000).

This new experiment evaluates the influence of the number of alternatives in the antecedent by using new contents but maintaining the rest of the conditions present in Experiment 1. Instead of providing conclusions that involve explicit negations ("it is not even"), it employs implicit ones ("it is odd"). Thus, a simple contrast class (even-odd) is regarded as equivalent to an explicit negation "it is not even = it is odd", whereas in the case of a higher class (a letter), the negation "it is not u" can lead participants to think about more alternatives (other vowels: "e", "i"...). Since MP inference was not informative for the evaluation of the hypotheses, this study includes only MT inference. Therefore, the new conditions will make it possible to detect whether or not responses are produced through a search for alternatives, which guarantees that correct responses are found only after the search. If there are variations in the way of processing attribute contents – if it is more difficult, as hypothesised in Experiment 1– differences should emerge as a benefit for participants with a higher WM span who use the search strategy (Prediction 3).

As in the previous study, a correlation between MT logical inferences and the WM test (RST) is expected (Prediction 2). In addition, there should be differences in the difficulty of attribute and between-objects contents and between participants only in those with high WM capacity (Prediction 3). With the traditional inference task, correct MT responses can be obtained when participants consider alternatives and when they do not, as happened in Experiment 1. The inclusion of the "Number of alternatives" variable in the manipulation of the contents of the relationship (attribute and between-objects), is intended to detect those participants who consider alternatives with MT inferences. With multiple alternatives, the only correct response is "there is no possible conclusion", rather than not-p. If the participants look for alternatives while inferring, as the mental model theory proposes, the response time for conditionals with multiple alternatives will have greater latencies than for single alternative contents. Also, the number of right answers will be greater for this type of conditional when WM span is high (Prediction 4). Again, we expect to obtain correlations between the frequency of correct responses to multiple-alternative conditionals and the WM test. The specific predictions of deductive theories on TM inferences according to working memory group and the effect on thinking according to the number of alternatives can be seen in Table 4.

Method

Participants

A total of 81 university students, 12 male and 69 female, were evaluated. Their average age was 20 years and 8 months (age range: from 18 years 4 months to 37 years 1 month). All participants belonged to the Spanish-speaking University of Granada and agreed to participate voluntarily through informed consent. The project was reviewed and approved by the ethics committee of the University of Granada.

Design and Materials

The conditional reasoning task was the same used in Experiment 1, but making some methodological changes. The accuracy of the participants' responses and the time taken was recorded (includes time of the statement, the premise and the answer). In addition, the RST was employed to measure

Table 4 Predictions of the deductive theories about MT inferences for High and Low working memory groups and the effect on thinking with single or multiple concrete alternatives to the antecedent

WM Groups	Effect predicted	Mental Rule theories	Suppositional theory	Mental model theory
Low	Number of alternatives	No	No	No
High	Number of alternatives	No	No	Yes Multiple harder than simple

WM capacity. The independent variables considered were: contents of the relationship (attribute and between-objects), number of alternatives (single and multiple) and WM group (high or low). The dependent variables considered were the accuracy and latency of the participants' answers. The general analysis utilised a 2x2x2 design with repeated measures ANOVA and factorial ANOVA for specific comparisons. The data were subjected to intra-subject and inter-group analyses.

Conditional Reasoning Task

As in the previous study, a computer version was used (E-Prime) to measure the accuracy and latency of the participants' responses. First, the instructions and a practice example appeared, followed by 22 arbitrary contents of factual conditional statements with their respective categorical premises and multiple answer options. They included 16 MT logical inferences, constituting the experimental tests, and 6 logical inferences (2 MP, 2 AC, and 2 NA) used as filler material (see Appendix 2). The contents of the conditional statements were distributed homogeneously, so as to include statements with between-objects relationships and attribute ones. In addition, the study included the "number of alternatives of the antecedent" variable, which considered single-alternative antecedents (S) and multiple-alternative ones (M). To verify the number of alternatives that could be generated by the contents, 13 participants were chosen for a pre-test, in which they were given a series of phrases and had to provide one or more alternatives. The parameter used to regard contents as involving one or multiple alternatives was that at least 75% responded in the same way. Both possibilities were distributed homogeneously in MT inferences, as were the contents of the relationship. Questions were asked randomly. Each participant was presented with only one version. They were all controlled to have the same number of syllables in their contents (Spanish version), and their positive and negative alternatives were shown in the same order (positive when it corresponded to the antecedent, and negative when it was opposed to or different from the antecedent). The alternatives were programmed so that, for a given content, a positive alternative appeared the first time, and a negative one the next time. In all cases, the negations were implicit (for example, "it is odd" rather than "it is not even").

Afterwards, an example of MT logical inferences was shown, considering the content of the relationship between the antecedent and the consequent and the number of alternatives of the antecedent:

Attribute-related contents with one alternative for the antecedent:

If the person is alive, then he is a man
He is not a man
Therefore,
1) The person is alive (positive alternative)
2) The person is dead (negative alternative)
3) It is unknown

Attribute-related contents with multiple alternatives for the antecedent:

If the toy is red, then it is a motorcycle
It is not a motorcycle
Therefore,
1) The toy is red (positive alternative)
2) The toy is brown (negative alternative)
3) It is unknown

Between-objects contents with one alternative of the antecedent:

If the horse is healthy, then there is a monkey.
There is not a monkey.

Therefore,

```
1) The horse is healthy (positive alternative)
2) The horse is sick (negative alternative)
3) It is unknown
```

Between-objects contents with multiple alternatives of the antecedent:

If the ring is made of silver, then there is a rifle
There is not a rifle
Therefore,

```
1) The ring is made of silver (positive alternative)
2) The ring is made of gold (negative alternative)
3) It is unknown
```

The instructions are similar to those of Experiment 1, except for a few variations in the examples, in accordance with the new manipulations.

Procedure

The two tasks (conditional reasoning and WM test) were completed by all participants in 25-min individual sessions, held in a room at the university. The order was controlled by counter-balance: (1) Conditional reasoning, (2) RST. The series, the way of correcting the WM test and the researcher's function were exactly the same as in Experiment 1.

Results and Discussion

Based on the scoring of RST and as the variable had a normal distribution, participants were divided into two WM groups according to quartile 1 low (≤ 2.8; 37 participants) and quartile 4 high (≥ 3.4; 25 participants), using the extreme groups design presented by Conway et al. (2005).

Two variance analyses (repeated measures ANOVA) were also conducted, one with accuracy values and the other with latencies, considering the following factors: Number of alternatives (single and multiple) x Content (Attribute and Between-objects) x WM Group (High or Low). Data were analysed using SPSS 22.0 software.

Analysis of Accuracy

The ANOVA showed no differences in main effect of the contents of the relationship, the number of alternatives or WM groups. The repeat measures ANOVA showed two interaction effects. The first was between the contents of the relationship (64% attribute and 60% between-objects) and the number of alternatives (61% single and 62% multiple) ($F(1, 60) = 6.163$, $p = 0.016$, partial $\eta 2 = 0.093$) in which the differences were found between attribute single and between-objects single ($F(1, 60) = 4.629$, $p = 0.034$, partial $\eta 2 = 0.055$) with more correct answers in attribute single. The second interaction effect was between the contents of the relationship, the number of alternatives and WM groups ($F(1, 60) = 4.103$, $p = 0.047$, partial $\eta 2 = 0.064$) (Table 5). As Fig. 1 shows, more accuracy was observed in the high WM group; this result was obtained with the following contents: attribute single alternative ($F(1, 61) = 4.647$, $p = 0.035$, partial $\eta 2 = 0.072$), attribute multiple alternatives ($F(1, 56) = 5.406$, $p = 0.023$, partial $\eta 2 = 0.083$), and between-objects multiple alternatives ($F(1, 61) = 5.280$, $p = 0.025$, partial $\eta 2 = 0.081$).

In addition, high and low WM groups were analysed separately. The analysis considered the correct answers to the conditional reasoning task according to the contents of the relationship (attribute and between-objects) and the number of alternatives of the antecedent (single and multiple). No differences were observed in the low WM group between the "content" and "number of alternatives" variables. In the high WM group, correct answers for attribute contents were more frequent than for between-objects contents (82% vs 74%; $F(1, 24) = 4.571$, p = 0,043, partial $\eta 2 = 0.160$). Only in the high WM group is the effect of the content consistent with Prediction 3). The inclusion of multiple alternatives allowed us to test Prediction 4: the search for alternatives is shown clearly only in the high WM group. An interaction effect was observed between "number of alternatives" and "content" ($F(1, 24) = 6.000$, $p = 0.022$, partial $\eta 2 = 0.200$), with a higher percentage of correct answers when multiple alternatives were presented (68% single alternative and 88% multiple alternatives); this difference appeared only in the case of between-objects contents ($F(1, 24) = 6.383$, $p = 0.019$, partial $\eta 2 = 0.210$). These results are consistent with the predicted effect of content (Prediction 3). Figure 1 displays the data obtained.

To choose single alternative contents correctly in MT inferences, participants had to deny the consequent, just as in Experiment 1. However, to choose multiple-alternative contents correctly, they had to answer with the "it is unknown" option. Therefore, the comparison of results associated with the two contents (with single and multiple alternatives) has to be performed cautiously.

Table 5 Percentage of MT answers, in general and in the two WM groups according to the contents of the relationship and the number of alternatives, Experiment 2: Correct answers are in bold

	Attribute Single alt.	Attribute Mult. alt.	Betw. Object Single alt.	Betw. Object Mult. alt.
General (*N*=81)				
Yes	18	21	21	18
No	**65**	17	**57**	20
It is unknown	17	**62**	22	**62**
Low WM (*N*=37)				
Yes	25	25	23	22
No	**56**	17	**56**	18
It is unknown	19	**58**	21	**60**
High WM (*N*=25)				
Yes	4	9	13	3
No	**78**	5	**58**	7
It is unknown	18	**86**	29	**90**

Correct answers are in bold (please note that the general score is the average of all groups, not only of high and low WM groups)

Analysis of Latency Responses

The latency analysis (with a logarithmic transform) revealed differences depending on the number of alternatives, with longer time values for multiple alternatives (8.38 single and 8.50 multiple) ($F(1, 44) = 12.722$, $p = 0.001$, partial $\eta 2 = 0.224$) (see Table 6). Correct responses in multiple alternatives require the search for alternatives, which takes time (Prediction 4). The interaction effect between contents (attribute and between-objects) and number of alternatives ($F(1, 44) = 4.274$, $p = 0.045$, partial $\eta 2 = 0.089$) showed that this difference between the number of alternatives and the tendency towards higher time values for multiple alternatives reappeared with attribute contents ($F(1, 44) = 3717$, $p = 0.032$, partial $\eta 2 = 0.145$) and between-objects contents ($F(1, 46) = 15.486$, $p = 0.000$, partial $\eta 2 = 0.252$).

Lastly, an interaction effect between WM capacity, number of alternatives and content was observed ($F(1, 44) = 4.274$, $p = 0.045$, partial $\eta 2 = 0.090$). The low WM group showed differences only in the "number of alternatives" variable (single and multiple) ($F(1, 24) = 11.572$, $p = 0.003$, partial $\eta 2 = 0.345$), with longer response times for multiple alternatives (8.53) than for a single alternative (8.39). In the high WM group, there were differences between attribute and between-

Fig. 1 Percentage of correct answers to MT inferences, according to the contents of the relationship and the number of alternatives, Experiment 2: Low WM (left) and high WM (right) groups

Table 6 Latency (Standard Deviation) of correct answers to MT inferences in low and high WM groups, according to the contents of the relationship (Attribute and Between-objects) and the number of alternatives (Single or Multiple), Experiment 2

	Att. S.	Att. M.	BObj. S.	BObj. M.
General	8.42(.28)	8.50(.31)	8.33(.24)	8.50(.28)
Low	8.39(.22)	8.53(.29)	8.32(.32)	8.55(.29)
High	8.46(.35)	8.47(.29)	8.35(.25)	8.42(.25)

objects contents ($F(1, 24) = 8007$, $p = 0{,}010$, partial $\eta 2 = 0.267$), with longer time values for the former (8.47 attribute and 8.39 between-objects). This replicates the results of Experiment 1. In addition, differences were observed in the interaction between the "content" variable (attribute and between-objects) and the number of alternatives provided (single and multiple) ($F(1, 39) = 10.795$, $p = 0.002$, partial $\eta 2 = 0.217$). Specifically, there were differences between single-alternative and multiple-alternative between-objects contents ($F(1, 44) = 7.004$, $p = 0.011$, partial $\eta 2 = 0.137$), with the latter displaying higher time values. This was not true for attribute contents with a single alternative or with multiple alternatives ($F(1, 43) = 3.841$, $p = 0.057$, partial $\eta 2 = 0.082$) . These results are consistent with the accuracy data. A difference was observed between single alternative attribute contents and single alternative between-objects contents ($F(1,17) = 6.280$, $p = 0.003$, partial $\eta 2 = 0.178$), with the former displaying higher time values.

Correlations

Table 7 shows the correlation between the percentage of correct answers in the conditional reasoning task (MT inference), the content of the relationship (attribute and between-objects), the number of alternatives to the antecedent (single and multiple) and the RST. The correlation between MT inferences and working memory is significant ($r = 0.28$, $p = 0.013$), but only when the number of alternatives to the antecedent is multiple (Attribute $r = 0.25$, $p = 0.024$ / between-objects $r = 0.26$, $p = 0.020$); that is, when the correct answer involves considering alternative antecedents.

In Experiment 2, the inclusion of implicit denials (instead of "it is not red", "it is white") allowed us to detect when the correct answers to MT inferences were the product of a prior search for alternatives. This procedure prevented the recording of non-discriminating answers as correct (for instance, "it is not red") when the person considered a single case ("it is blue") while correctly assessing the answers of participants who considered different alternatives (blue, brown, white...). In the case of implicit denials with a single alternative (for example, "on", "off"), the condition was similar to the explicit denial used in Experiment 1 ("white" or "not white"). In both cases, denying a category implied affirming the opposite one. The advantage of attribute-related contents followed the tendency of Experiment 1. The low difficulty of single alternative attribute contents can be initially explained because in this case the antecedent and the consequent had a direction, which facilitated participants' access to contents, which allowed them to generate alternatives more easily. The reason the difference was not present between attribute-related and between-objects contents when multiple alternatives were provided is that the multiple alternatives of the antecedent led individuals to think about different contents, which facilitated access to alternatives in long-term memory, independently of the contents used.

The analysis of contents (attribute and between-objects) and the number of alternatives (single and multiple) revealed latency differences, but no variations in accuracy between attribute and between-objects multiple. The differences in latency were predictable, as the manipulation of the number of alternatives should facilitate fleshing out alternative models, and therefore should result in response time differences. The data revealed higher time values for contents with multiple alternatives compared with single alternative contents. These differences can be explained, as people dealing with multiple-alternative MT inferences may find it easier to access or flesh out other mental representations. The results are clearer when considering high and low WM groups. Higher response times appeared in answers to multiple-alternative contents, both with high and low WM capacities, but the high WM group was significantly more successful. Although low WM individuals tried to access alternative models, and thus increased their response times, they did not manage to improve the accuracy of their responses. In contrast, the greater processing space available to high-WM participants allowed them to increase their number of correct answers. In other words, successfully searching for alternatives, which involves the processing and manipulation of representations, requires a high WM capacity as well as access to contents.

Table 7 Correlation between WM capacity and correct answers to MT inferences with different contents, Experiment 2

		RST	
MT	0.28**		
Attribute		0.29*	
Single alt.			0.21#
Multiple alt.			0.25*
Between-objects		0.24*	
Single alt.			0.07
Multiple alt.			0.26*

* $p < 0.05$ ** $p < 0.001$ # $p = 0.06$

The correlation data point in the same direction. The frequency of correct answers in MT inferences was positively correlated with the RST, particularly in the multiple condition.

A noteworthy aspect is that the differences due to the influence of attribute-related and between-objects contents could only be observed in the high WM group. As in the previous study, high WM participants took longer to solve tasks with attribute contents and single alternative (when the correct response was "not p" as in the previous study); in this case, though, the new task did reveal their greater accuracy in the same conditions (but with between-objects contents), and also compared with the low WM group. In the case of between-objects contents and a single alternative, as in Experiment 1, there were no differences between high and low WM groups. This was the only condition in which no differences were observable.

General Discussion and Conclusions

Previous results have shown: 1. There is a general relationship between WM capacity and reasoning performance, explained by the assumption that reasoning requires an analytic thinking process that taps WM resources (see, Verschueren et al., 2005). 2. Reasoning problems are often solved by heuristic processes, which do not rely as heavily on WM capacity (Evans, 1993; Klaczynski, 2001). Research on syllogistic reasoning has shown that highly skilled participants used more demanding reasoning processes while reasoners with more limited WM capacity relied on less demanding processes (Aslan & Bäuml, 2010; Braver et al., 2007; Copeland & Radvansky, 2007). These facts could explain some of the present results.

Using an heuristic strategy to solve valid inferences (MP and MT) is open to all participants, but using a more analytical strategy that requires more complex resources (such as *reductio ad absurdum* or the search for alternatives) will be available for participants with enough resources and therefore is more likely to be used by high working memory participants. In order to maximise differences and to exclude participants with intermediate WM capacity who might be using different strategies, and following Conway et al. (2005) we selected only the participants in high and low working memory groups. The results reveal the advantage of the more automatic inference, MP, compared with MT: shorter response times and a greater number of correct answers. The most relevant manipulation in this study is the connection of the conditional relationship between the antecedent and the consequent with the alternatives to the antecedent. Both factors have been shown to be clearly associated with a person's WM. Although the more automatic MP inferences were not affected or associated with WM span, either in the accuracy of the participants' answers or the time that they spent inferring, the frequency of correct MT inferences was related to WM span. This correlation, observed in both experiments, is consistent with the results of Toms et al. (1993).

Curiously, although in Experiment 1, high WM participants did not give more correct answers to MT inferences, they spent more time inferring when the contents of the conditional established an attribute relationship between the antecedent and the consequent. In Experiment 1, two different strategies to solve the MT inferences could lead to the same correct answer (if p then q; not q therefore, not p): one is looking for alternatives to p (p1,p2…pn), and, the other, for a faster strategy, just discarding p (not p). The first one takes longer and could be used by higher WM participants. Actually, in Experiment 2 when the only correct response required was to look for alternatives (asking for p1 or pn), the higher WM participants gave more correct responses.

The distinction between automatic and controlled processing in deduction, with Type 1 processing and Type 2 processing respectively, could help us to understand the present results. Oaksford and Chater (2010) proposed that the activation and recovery of associated representations in LTM is an automatic process based on Type 1. However, they proposed that Type 2 is required for the storage and manipulation of the information given by the premises and is completed with the information retrieved from LTM. Therefore, Type 2 (and the WM) is required particularly when multiple alternatives have to be considered. This is what seems to happen in Experiment 2. Only those participants with sufficient WM resources available can use the controlled strategy. Verschueren et al. (2005) assumed that those participants with high WM capacity prefer to reason analytically, because conclusions based on counterexample information are considered to be more informative and accurate than those based on likelihood estimates. As in the present study, Verschueren et al. (2005), using everyday causal sentences, found that reasoners with a lower WM span relied more often on a simple probabilistic reasoning mechanism and only highly skilled reasoners took counterexample information into account in the way described by the mental model theory.

The hypothesis that representational structure influences long-term memory, which some authors have highlighted as relevant for making inferences (Barrouillet, 2011; Barrouillet & Lecas, 2002; Markovits, 2000; Oaksford & Chater, 1994), seemed to be confirmed, but only in the case of high WM individuals. In attribute relationships, a single object appears to be represented: "if there is a triangle, then it is red". In contrast, two objects are represented in between-objects relationships: "if there is a triangle, then there is a square." MT inference involves discarding the consequent ("it is not red" and "there is not a square", respectively). Understanding the premise should lead to discarding the objects mentioned, but while the attribute-related content leaves us without a triangle, the between-objects content leaves one representation of a triangle which we must discard. Although the inference may result in more errors in the second case (between-objects, since the representation of the triangle remains), the time available to look for alternatives (no triangle) should be more in the first case, since we are left without referents.

Therefore, as Verschueren et al. (2005) found and as this study also shows, it is possible that the strategies used by the participants, depending on their WM resources, may have led them to use a more automatic strategy, or one focused on the search for alternatives. The task in Experiment 1, with explicit denials, could not display the benefits of searching for alternatives, but the search itself was manifested through the longer period spent reaching the inference. In Experiment 2, which included implicit denials, correct answers could only be found by considering different alternatives. The replication of this experiment again showed the time difference, but with an added improvement in accuracy. The results were consistent with the differential influence of the type of relationship involved, although only in the case of participants with a high WM capacity.

Barrouillet and Lecas (2002, see also Barrouillet, 2011) have justified the importance of the access of memory to the representational semantic network during inference. One of their arguments is based on the different frequency of the inferences made depending on whether the contents of the conditional were binary or involved multiple alternatives. Through different tasks, researchers have studied inferences to compare the search for alternatives when single alternative or binary contents (day-night) are provided and when multiple-alternative contents are present (Monday-Tuesday-Wednesday; for example, see Oaksford, 2002; Oaksford & Chater, 1994; Schroyens et al., 2000). Implicit denial requires that the individuals look for alternatives depending on those provided by the content. In Experiment 2, we added the same manipulation (Barrouillet & Lecas, 2002) of the number of alternatives: single (which they refer to as "binary") or multiple. The manipulation was effective, and showed differences in the ease of accessing single or multiple alternatives, but these differences were only observed in participants with a high WM capacity. Multiple-alternative contents displayed higher latencies than single-alternative contents. The main characteristic of this manipulation was that the high WM group was more precise. This result can be construed as suggesting that searching for other alternative models efficiently, that is, the ability to manipulate models or representations, requires a high WM capacity. Consistent with the above, it is relevant that the correlation of MT logical inferences with the test to measure WM capacity (RST) was observed with all contents (attribute with a single alternative and with multiple alternatives, and between-objects with a single alternative), except for between-objects content with a single alternative.

We believe that the results are relevant in several ways. First, the results suggest that low and high WM participants used different strategies. Possibly the first group used a more automatic strategy (based on Type 1 processing) and the second, the more controlled (Type 2 processing). In any case, the participants all applied processes to solve MP inferences more quickly than MT ones, and with a higher success rate. Contents and the number of alternatives to the antecedent only had an effect on the inferences of the high WM group. Second, although the results can be interpreted on the basis of various models and theories of conditional reasoning, they provide restrictions that can help us to refine our predictions about inference. For example, the results obtained are consistent with what Barrouillet and Lecas (2002) have proposed, but with a variation: the possible influence of the knowledge structure of long-term memory would be mediated by WM capacity, that is to say, by individuals' capacity to access it and possibly by the strategy that they use.

We have seen that deductive theories make very similar predictions (see Tables 1 and 4). However, those participants with high working memory capacity, who could apply a controlled strategy, fit well with predictions of the mental model theory. This theory, unlike the other deductive theories, establishes that the structure of the mental representations has a role in the inference process. Halford et al.'s (2010) relational complexity proposal also capture the importance of the relation in the representation. They proposed that the working memory is a workspace where relational representations are constructed. The kind of relation established between the entities (antecedent and consequent) is transported into mental models. The complexity operating with those models, might depend on the nature of the relation and on the capacity of working memory. When participants have to mentally manipulate with one ore multiple attributes of an object, the relationships are formed in a different way than when the relationship is just a compound of objects.

Finally, there are some limitations of the present research. Participants were all undergraduate students; as in many other studies in psychology, the extent to which they represent the general population is limited. Also, we decided to use the RST test of the WM verbal component. Other WM tests (such as the Digit span test) and also other cognitive measures such as the need for cognition would help to obtain additional and more detailed information in this research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12144-021-01966-3.

Acknowledgments We would like to highlight the invaluable contributions of Orlando Espino, Ana Cristina Quelhas, Antonio Rus and Carlos Santamaría.

Funding This study was financed by the Chilean government's National Fund for Scientific and Technological Development (projects FONDECYT-ANID number 1210989), and by the Spanish Government, Ministry of Economy and Competitiveness (PGC2018-095868-B-I00).

This study was supported by the Associative Research Program (PIA) on Cognitive Sciences, Research Center on Cognitive Sciences (CICC), Faculty of Psychology. Universidad de Talca, Campus Lircay, Chile.

Data Availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical Approval The study protocol was approved by the ethics committee of University of Granada.

Informed Consent Informed consent was given to all participants in order to get their allowance for this study.

References

Alloway, T. P., & Copello, E. (2013). Working memory: The what, the why, and the how. *The Australian Educational and Developmental Psychologist, 30*(2), 105–118.

Anderson, J. R., & Lebiere, C. (1998). *Atomic components of thought*. Erlbaum.

Aslan, A., & Bäuml, K.-H. (2010). Retrieval-induced forgetting in young children. *Psychonomic Bulletin & Review, 17*, 704–709. https://doi.org/10.3758/PBR.17.5.704.

Baddeley, A. D. (1986). *Working memory*. Oxford University Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2.

Baddeley, A. D. (2007). *Working memory, thought and action*. Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (pp. 47–89). Academic Press.

Baddeley, A. D., Eysenck, M. W., & Anderson, M. (2009). *Memory*. Psychology Press.

Barrouillet, P. (2011). Dual-process theories of reasoning: The test of development. *Developmental Review, 31*(2), 151–179. https://doi.org/10.1016/j.dr.2011.07.006.

Barrouillet, P., & Lecas, J. F. (1999). Mental models in conditional reasoning and working memory. *Thinking and Reasoning, 5*, 289–302. https://doi.org/10.1080/135467899393940.

Barrouillet, P., & Lecas, J. F. (2002). Content and context effects in children's and adults' conditional reasoning. *Quarterly Journal of Experimental Psychology-A, 55*(3), 839–854. https://doi.org/10.1080/02724980143000587.

Barrouillet, P., Gauffroy, C., & Lecas, J. F. (2008). Mental models and the suppositional account of conditionals. *Psychological Review, 115*(3), 760–771. https://doi.org/10.1037/0033-295X.115.3.760.

Bjorklund, D. (2005). *Children's thinking: Cognitive development and individual differences*. Wadsworth/Thomson.

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review, 85*, 1–21. https://doi.org/10.1037/0033-295x.85.1.1.

Braine, M. D. S., & O'Brein, D. P. (1991). A theory of if: A lexical entry, reasoning program and pragmatic principles. *Psychological Review, 98*, 182–203. https://doi.org/10.1037//0033-295X.98.2.182.

Braine, M. D. S., & O'Brein, D. P. (Eds.). (1998). *Mental logic*. Lawrence Erlbaum Associates.

Braine, M. D. S., Reiser, B. J., & Rumain, B. (1984). Some empirical justification for a theory of mental propositional logic. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*. Academic Press.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 76–106). Oxford University Press.

Cantor, J., & Engle, R. W. (1993). Working-memory capacity as long-term memory activation: An individual-differences approach. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 5*, 1101–1114. https://doi.org/10.1037//0278-7393.19.5.1101.

Case, R. (1981). Intellectual development: A systematic reinterpretation. In F. H. Farley & N. J. Gordon (Eds.), *Psychology and education*. McCutchan Publishing Corporation.

Case, R. (1985). *Intellectual development. Birth to adulthood*. Academic Press.

Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology, 38*, 191–258. https://doi.org/10.1006/cogp.1998.0696.

Conway, A. R. A., Kane, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786. https://doi.org/10.3758/BF03196772.

Copeland, D. E., & Radvansky, G. A. (2007). Aging and integrating spatial mental models. *Psychology and Aging, 22*(3), 569–579. https://doi.org/10.1037/0882-7974.22.3.569.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–185. https://doi.org/10.1017/S0140525X01003922.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6.

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science, 17*, 428–433. https://doi.org/10.1111/j.1467-9280.2006.01723.x.

De Neys, W., Schaeken, W., & D'Ydeawalle, G. (2005). Working memory and counterexample retrieval for causal conditionals. *Thinking and Reasoning, 11*(2), 123–150. https://doi.org/10.1080/13546780442000222.

Elosúa, M. R., Gutiérrez, F., García-Madruga, J. A., Luque, J. L., & Gárate, M. (1996). Adaptación española del Reading Span Test de Daneman y Carpenter. *Psicothema, 8*(2), 383–395.

Engle, R. W., & Oransky, N. (1999). The evolution from short-term to working memory: Multi-store to dynamic models of temporary storage. In R. Sternberg (Ed.), *The nature of human cognition*. MIT Press.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and function of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge University Press.

Evans, J. S. B. T. (1993). The mental model theory of conditional reasoning: Critical appraisal and revision. *Cognition, 48*, 1–20. https://doi.org/10.1016/0010-0277(93)90056-2.

Evans, J. S. B. T., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking and Reasoning, 11*, 382–389. https://doi.org/10.1080/13546780542000005.

Evans, J. S. B. T., & Handley, S. J. (1999). The role of negation in conditional inference. *Quarterly Journal of Experimental Psychology-A, 52*(3), 39–769. https://doi.org/10.1080/713755834.

Evans, J. S. B. T., & Over, D. E. (2004). *If*. Oxford University Press.

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*(3), 223–241. https://doi.org/10.1177/1745691612460685.

Evans, J. S. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology - Learning, Memory, and Cognition, 29*(2), 321–355. https://doi.org/10.1037/0278-7393.29.2.321.

García-Madruga, J. A., Gutiérrez, F., Carriedo, N., Luzón, J. M., & Vila, J. O. (2007). Mental models in propositional reasoning and working memory's central executive. *Thinking and Reasoning, 13*(4), 370–393. https://doi.org/10.1080/13546780701203813.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177.

Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences, 14*(11), 497–505. https://doi.org/10.1016/j.tics.2010.08.005.

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

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

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109*(4), 646–678. https://doi.org/10.1037/0033-295X.109.4.646.

Khemlania, S., & Johnson-Laird, P. N. (2012). The processes of inference. *Argument and Computation, 4*, 1–17. https://doi.org/10.1080/19462166.2012.674060.

Klaczynski, P. A. (2001). Analytic and herusitic processing influences on adolescent reasoning and decision making. *Child Development, 72*(3), 844–861. https://doi.org/10.1111/1467-8624.00319.

Klauer, K. C., Stegmaier, R., & Meiser, T. (1997). Working memory involvement in propositional and spatial reasoning. *Thinking and Reasoning, 3*, 9–48. https://doi.org/10.1080/135467897394419.

Kruglanski, A. W. (2013). Only one? The default interventionist perspective as a Unimodel. Commentary on Evans & Stanovich. *Perspectives on Psychological Science, 8*(3), 242–247. https://doi.org/10.1177/1745691613483477.

Markovits, H. (2000). A mental model analysis of young children's conditional reasoning with meaningful premises. *Thinking and Reasoning, 6*, 335–347. https://doi.org/10.1080/135467800750038166.

Markovits, H. (2014). Conditional reasoning and semantic memory retrieval. In A. Feeney & V. Thompson (Eds.), *Reasoning and memory*. Psychology Press.

Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review, 22*, 5–36. https://doi.org/10.1006/drev.2000.0533.

Markovits, H., Doyon, C., & Simoneau, M. (2002). Individual differences in working memory and conditional reasoning with concrete and abstract content. *Thinking and Reasoning, 8*, 97–107. https://doi.org/10.1080/13546780143000143.

Oaksford, M. (2002). Contrast classes and matching bias as explanations of the effects of negations on conditional reasoning. *Thinking and Reasoning, 8*, 135–151. https://doi.org/10.1080/13546780143000170.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*, 608–631. https://doi.org/10.1037/0033-295X.101.4.608.

Oaksford, M., & Chater, N. (2010). Conditional inference and constraint satisfaction: reconciling mental models and the probabilistic approach? In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 309–334). Oxford, UK: Oxford University Press.

Psychology Software Tools (2007). *E-Prime v2.0 (Release Candidate)* [Computer software]. Psychology Software Tools.

Ricco, R. B., Koshino, H., Sierra, A. N., Bonsel, J., Von Monteza, J., & Owens, D. (2020). Individual differences in analytical thinking and complexity of inference in conditional reasoning. *Thinking & Reasoning*., 1–31. https://doi.org/10.1080/13546783.2020.1794958.

Rips, L. J. (1994). *The psychology of proof*. Routledge.

Schaeken, W., & Schroyens, W. (2000). The effect of explicit negations and of different contrastclasses on conditional syllogisms. *British Journal of Psychology, 91*, 533–550.

Schroyens, W., Schaeken, W., Verschueren, N., & D'Ydewalle, G. (2000). Conditional reasoning with negations: Implicit and explicit affirmation or denial and the role of contrast classes. *Thinking and Reasoning, 6*, 221–252. https://doi.org/10.1080/13546780050114519.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Blackwell.

Stanovich, K. E. (2011). *Rationality and the reflective mind*. Oxford University Press.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94*(4), 672–695. https://doi.org/10.1037/0022-3514.94.4.672.

Toms, M., Morris, N., & Ward, D. (1993). Working memory and conditional reasoning. *Quarterly Journal of Experimental Psychology-A, 46*, 679–699. https://doi.org/10.1080/14640749308401033.

Vandierendonck, A., Dierckx, V., & van der Beken, H. (2006). Interaction of knowledge and working memory in reasoning about relations. In C. Held, M. Knauff, & G. Vosgerau (Eds.), *Mental models and the mind: Current developments in cognitive psychology, neuroscience, and philosophy of mind*. Elsevier.

Verschueren, N., Schaeken, W., & D'Ydewarlle, G. (2005). Everyday conditional reasoning: A working memory-dependent tradeoff between counterexample and likelihood use. *Memory & Cognition, 33*(1), 107–119. https://doi.org/10.3758/BF03195301.

Whitney, P., Arnett, P., Driver, A., & Budd, D. (2001). Measuring central executive functioning: What's in a reading span? *Brain and Cognition, 45*, 1–14. https://doi.org/10.1006/brcg.2000.1243.