



What we count dictates how we count: A tale of two encodings

Hippolyte Gros, Jean-Pierre Thibaut, Emmanuel Sander

► To cite this version:

Hippolyte Gros, Jean-Pierre Thibaut, Emmanuel Sander. What we count dictates how we count: A tale of two encodings. *Cognition*, 2021, 212, pp.104665. 10.1016/j.cognition.2021.104665. hal-03198508

HAL Id: hal-03198508

<https://hal.science/hal-03198508>

Submitted on 24 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

What we count dictates how we count: A tale of two encodings

Hippolyte Gros^{1,*} · Jean-Pierre Thibaut² · Emmanuel Sander¹

¹University of Geneva · ²University of Bourgogne Franche-Comté

Abstract:

We argue that *what* we count has a crucial impact on *how* we count, to the extent that even adults may have difficulty using elementary mathematical notions in concrete situations. Specifically, we investigate how the use of certain types of quantities (durations, heights, number of floors) may emphasize the ordinality of the numbers featured in a problem, whereas other quantities (collections, weights, prices) may emphasize the cardinality of the depicted numerical situations. We suggest that this distinction leads to the construction of one of two possible encodings, either a cardinal or an ordinal representation. This difference should, in turn, constrain the way we approach problems, influencing our mathematical reasoning in multiple activities. This hypothesis is tested in six experiments ($N = 916$), using different versions of multiple-strategy arithmetic word problems. We show that the distinction between cardinal and ordinal quantities predicts problem sorting (Experiment 1), perception of similarity between problems (Experiment 2), direct problem comparison (Experiment 3), choice of a solving algorithm (Experiment 4), problem solvability estimation (Experiment 5) and solution validity assessment (Experiment 6). The results provide converging clues shedding light into the fundamental importance of the cardinal versus ordinal distinction on adults' reasoning about numerical situations. Overall, we report multiple evidence that general, non-mathematical knowledge associated with the use of different quantities shapes adults' encoding, recoding and solving of mathematical word problems. The implications regarding mathematical cognition and theories of arithmetic problem solving are discussed.

Author Note

Data and materials for all experiments are available online at https://osf.io/kz6gh/?view_only=2f3fb6b910844e238ae58fef3c61168a. This research was supported by grants from the Regional Council of Burgundy, Paris Feder Grants (20159201AAO050S02982 & 20169201AAO050S01845), the Experimental Fund for the Youth and French Ministry of Education (HAP10-CRE-EXPE-S1), and from the French Ministry of Education and Future Investment Plan (CS-032-15-836-ARITHM-0). We thank Jay Boussier and Katarina Gvozdic for helpful comments on an earlier draft of the manuscript. Parts of this work were presented at the 38th Annual meeting of the Cognitive Science Society (Experiments 5 and 6). Correspondence concerning this article should be addressed to Hippolyte Gros, IDEA Lab, University of Geneva, Geneva, 1205, Switzerland. E-mail: hippolyte.gros@unige.ch.

Introduction

Mathematical skills are frequently used in everyday life (Northcote & Marshall, 2016). Whether someone is calculating a bus fare, estimating the duration of the ride, deciding whether to go up the stairs or take the elevator, assessing if the added weight of everyone in the elevator is below the weight limit, counting how many cookies to buy for the coworkers, or evaluating at which height to put the cookie jar so that the kids won't reach it, people use mathematics in many aspects of their lives. However, being able to perform those quick calculations on a daily basis requires the ability to use mathematical knowledge in very tangible situations (Schoenfeld, 2009). Mastering such a skill is the focus of a long learning process taking place in school, notably involving mathematical word problems (Daroczy, Wolska, Meurers, & Nuerk, 2015; Stacey, 2005; Verschaffel, Schukajlow, Star, & Van Dooren, 2020). But while using arithmetic knowledge in daily-life might seem effortless to most adults, we argue that some underlying pervasive influence remains: in fact, even when using the most elementary mathematical skills, we have reasons to believe that *what* we count has a deep impact on *how* we count. This entails that even adults may be significantly better at using certain fundamental mathematical skills in situations involving bus rides, elevator trips or cookie jar heights than in situations pertaining to bus ticket prices, elevator weight limits or cookie counting. In this paper, we argue that such concrete situations evoke one of two distinct conceptions of numerical situations; they either emphasize the *cardinal* property of numbers or their *ordinal* property (see below). This difference leads to one of two possible encodings, which in turn constrain the way we represent word problems, influencing our mathematical reasoning in multiple activities.

An ontological distinction between cardinal and ordinal situations

This distinction between cardinality and ordinality is fundamental in mathematics, especially in set theory (Dantzig, 1945; Frege, 1980; Russell, 1919; Simon, 1997; Suppes, 1972). In common usage, ordinal numbers describe the numerical position of an object in an ordered sequence (i.e. 1st, 2nd, 3rd, etc.), whereas cardinal numbers refer to the general concept of quantity by designating the total number of entities within a set (Fuson, 1988; Wasner, Moeller, Fischer, & Nuerk, 2015). But this distinction has far-reaching implications beyond set theory and mathematics and, we believe, influences the very way humans comprehend numerosity in the world.

According to Piaget (1941/1965), the relationship between cardinality and ordinality is central to the notion of number (Fuson, 1988). Indeed, the two ideas are necessarily intertwined, since the cardinal and ordinal meanings of numbers are two sides of the same coin, two properties intrinsically tied to our use of numbers. As Fuson (1988, p. 363) puts it, “any ordinal number refers to the particular entity within a linear ordering that is

preceded by a cardinal number one less than it”. However, this interdependency between ordinality and cardinality does not preclude that a crucial difference between conceiving of numbers as cardinal count values or as order labels could exist.

From a developmental perspective, the seminal work of Gelman and Gallistel (1978) on counting principles introduced the idea that the understanding of the cardinal and ordinal properties of numbers could develop separately in children. In their work, Gelman and Gallistel proposed that **three** counting principles need to be mastered by children in their efforts to become proficient counters. **First**, the “one-to-one correspondence principle”, describes the fact that each of the items to be counted needs to be assigned one, and only one, counting word. Then, the “stable-order principle” is a necessary step in the development of ordinality, since it entails that children must learn that the list of words used to count needs to be said in a definite and stable order, each word having the same predecessor and the same successor over trials. Finally, the “cardinal principle” states that, provided that the one-to-one and the stable-order principles have been respected, the counting word allocated to the ultimate item in a collection represents the total number of items in that collection. Put together, these three principles called the “how-to-count” principles make it possible to understand the cardinality of numbers. Ever since this work, investigations have been conducted to study how children learn to master both the cardinal and the ordinal properties of numbers.

Children’s ability to comprehend and use the notion of cardinality in counting was thus scrutinized by numerous studies who showed its slow and sequential development over the first years of life (e.g. Bermejo, 1996; Condry & Spelke, 2008; Le Corre & Carey, 2007; Sarnecka & Lee, 2009; Wynn, 1992). As for the development of the ordinal meaning of numbers, it has been suggested that children do not learn to use ordinal labels such as “first” and “second” before they are 4 or 5 years old (Fischer & Beckey, 1990; Miller, Major, Shu, & Zhang, 2000; Miller, Marcovitch, Boseovski, & Lewkowicz, 2015). In recent years, the comparison of the developmental trajectories of the sense of cardinality and the sense of ordinality has received increasing research interest, with converging results showing that children tend to use cardinality before succeeding in ordinality tasks (Colomé & Noël, 2012; Meyer, Barbiers, & Weerman, 2016; Wasner et al., 2015). Further evidence for the importance of this distinction can be found in the brain; Delazer and Butterworth (1997) reported that a patient who suffered a cerebral lesion was left with impaired access to the cardinal meaning of numbers, but a selectively preserved access to their ordinal meaning. On the other hand, Turconi and Seron (2002) reported that a patient suffering from lesions in the right parietal occipital junction and in the left posterior parietal cortex struggled to judge ordinal relationships between numbers whereas he remained able to understand cardinality to a certain extent, as

shown by his successful completion of number comparison tasks. A similar dissociation was found by Chen, Xu, Shang, Peng, and Luo (2014) who reported that lesions in the left parietal lobe resulted in a selective impairment of order meaning among one of their patients. This dissociation between the two meanings of numbers was further supported by EEG studies who found that event-related potentials differed both in terms of timing and topography between order and quantity tasks (Rubinsten, Dana, Lavro, & Berger, 2013; Turconi, Jemel, Rossion, & Seron, 2004). Furthermore, an fMRI study by Lyons and Beilock (2013) showed that while cardinal and ordinal judgments tasks involved a similar frontoparietal network for non-symbolic quantities (i.e. dot arrays), the use of symbolic numbers (i.e. numerals) led to the activation of different networks for cardinal and ordinal tasks. Similarly, Matejko, Hutchison, and Ansari (2019) found evidence of ordinal processing in the left inferior parietal cortex of adults, while no such evidence was found for cardinal processing.

In recent years, behavioral studies have used distinct ordinality and cardinality tasks to investigate the orthogonality of these two sides of number processing. Using digit ordering and number comparison tasks, they showed that basic numerical measures of ordinal and cardinal processing differentially predict arithmetic performance and math achievement scores (Goffin & Ansari, 2016; Sasanguie, Lyons, De Smedt, & Reynvoet, 2017). Overall, these studies show the importance of the distinction between cardinality and ordinality in numerical cognition. Processing the cardinality and ordinality of numerical symbols involves distinct neural underpinnings, each equally important and each predicting independent variance in arithmetic and mathematical proficiency. As Lyons, Vogel, and Ansari put it (2016), while research has initially focused on cardinal processing, “overlooking the contribution of ordinality to how we process numbers is likely a major oversight” (p. 206).

Despite this growing body of research on cardinal and ordinal tasks, little is known about adults’ ability to perceive either the cardinality or the ordinality of complex numerical situations evoking those two sides of numbers. Can situations sharing the same mathematical deep structure highlight either cardinality or ordinality, depending on the daily-life knowledge they are imbued with? This paper aims at answering this crucial issue, by focusing on adults’ ability to use basic mathematical skills in situations meant to emphasize either the cardinal or the ordinal aspect of numbers. Indeed, we propose to investigate whether the presence of different quantities (e.g. weights versus durations) in the wording of otherwise mathematically identical problems may lead even adults to build either a cardinal or an ordinal encoding of the depicted situations. We thus scrutinize the claim that this distinction will, in turn, lead to clear-cut differences in the way these problems are categorized, compared, and solved.

The role of semantic knowledge in problem representation

Our main hypothesis is that the distinction between cardinal quantities and ordinal quantities will lead to clear-cut differences in the way arithmetic word problems are categorized, compared, and solved. This claim builds upon the literature on content effects in arithmetic word problems, which suggests that mathematically irrelevant semantic information may induce specific problem representations and that may interfere with arithmetic word problem solving. In 1995, Bassok, Wu, and Olseth showed that semantic knowledge linked to the different entities featured in a problem statement influenced analogical transfer between problems. They taught participants the algorithmic solution of a problem whose cover story depicted either objects assigned to people (e.g. computers given to secretaries) or people assigned to other people (e.g. doctors from one hospital assigned to doctors from another hospital). They then evaluated how well participants performed on problems sharing the same *objective mathematical structure* (i.e. what we call the problems' deep structure). Depending on whether the entities in the transfer problem had typically symmetrical roles (people and people) or typically asymmetrical roles (objects and people), participants' performance varied. Those results suggested that variations of semantic knowledge lead to different representations being abstracted. These *interpreted structures* either facilitated or hindered transfer depending on how well they mapped onto the problems' mathematical structure. Similarly, Bassok, Chase, and Martin (1998) asked participants to create addition or division word problems using specific sets of entities, either linked by a functional semantic relation (e.g. the container/content relation between fruit baskets and oranges) or by a collateral relation (e.g. oranges and apples belonging to the same superordinate "fruit" category). They showed that participants tended to propose division problems when the semantic knowledge induced by the entities evoked a functional relation, whereas they created addition problems when the entities were different kinds of fruits or other collateral elements.

Thus, the semantic knowledge induced by the entities described in the problem statement seems to influence the representation that participants encode. Bassok (2001) theorized this process by stating that an interpreted mathematical structure is abstracted by the participants, based on the semantic relations depicted in the problem statements (e.g. container/content or assigned/receiver relations). This interpreted structure can be semantically aligned or misaligned with the objective mathematical structure of the problem, depending on whether the two structures can be mapped onto each other (Bassok, 2001). For instance, the interpreted structure of a problem involving oranges and baskets will be semantically aligned with division, whereas the interpreted structure of a problem involving oranges and apples will be semantically aligned with addition instead. In support of this view, Bassok, Pedigo and Oskarsson's work (2008) on the priming of addition facts by different pairs of words is

enlightening. They showed that categorically related words (e.g. tigers and cheetahs) prime addition facts, whereas unrelated words (e.g. lungs and statues), as well as functionally related words (e.g. bears and claws), do not exert such priming. In other words, pairs of words semantically aligned with addition elicit an automatic retrieval of addition facts, whereas words misaligned with the structure of addition do not. Taken together, those studies suggest that the semantics induced by the entities manipulated in the problem statements influence the representational processes at play: non-mathematical properties of said entities constrain the operations selected by the participants.

Preliminary works on arithmetic word problem representations

In this paper, we push this point several steps further in the context of the cardinal-ordinal distinction. We investigate, with adults, the claim that the semantic knowledge about the entities described in a problem can evoke either an ordinal or a cardinal representation of the described situation, subsequently shaping mathematical reasoning. Regarding this matter, a study by Verschaffel, De Corte, and Vierstraete (1999) showed that upper elementary school pupils made specific errors when attempting to solve problems with ordinal numbers that they did not make on problems involving only cardinal numbers, thus suggesting that ordinality may interfere with arithmetic word problem solving processes. More recently, Gamo, Sander, and Richard (2010) undertook a study focusing on the perception of cardinality and ordinality in arithmetic word problems among 4th and 5th graders. The present paper builds upon Gamo et al.'s study, in which the authors created multiple-solution word problems sharing the same mathematical structure but differing in the type of quantities they used. Their problem statements revolved around, respectively, the number of family members in a hotel (family problems), the price of a series of items (price problems) or the age of a protagonist (age problems) (see Table 1). Gamo et al. hypothesized that family problems and price problems would both lead to a cardinal encoding of the situation described, whereas age problems would lead to an ordinal encoding of the situation. Consider the family problem (Table 1, left column). Gamo et al. showed that most participants solved it using a 3-step algorithm: $9 - 5 = 4$; $5 - 3 = 2$; $4 + 2 = 6$. Note, however, that this problem can also be solved with a one-step algorithm: $9 - 3 = 6$. Using this algorithm requires the participant to realize that since the Roberts are present at the hotel during both vacations, then the difference between the number of people in the Richards' and in the Dumas' families is equal to the difference between the total number of people at the hotel. Thus, calculating the number of people in the Roberts' family ($9 - 5 = 4$) or the number of people in the Dumas' family ($5 - 3 = 2$) is not necessary to find the solution.

Table 1: Family, price and age problems used in Gamo et al. (2010).

Family problem	Price problem	Age problem
In the Richards' family, there are 5 persons. When the Richards go on vacation with the Roberts, there are 9 persons at the hotel. In the Dumas' family, there are 3 people less than in the Richards' family. The Roberts go on vacation with the Dumas. How many will they be at the hotel?	John bought an 8-Euro exercise book and a pair of scissors. He paid 14 Euros. A pen costs 3 Euros less than the exercise book. Paul bought scissors and a pen. How much did he pay?	Antoine attended painting classes at the art school for 8 years and stopped when he was 17 years old. Jean began at the same age as Antoine and attended the course for 2 years less. At what age did Jean stop attending the classes?

Similarly, most participants use a 3-step algorithm to solve the price problem (Table 1, middle column). This algorithm consists in calculating the price of the pen and the price of the scissors, and then adding them up: $14 - 8 = 6$; $8 - 3 = 5$; $6 + 5 = 11$. But a 1-step algorithm can also be used to solve this problem: $14 - 3 = 11$. However, only participants who notice that John and Paul both bought scissors may also understand that the difference between the price of the exercise book and the price of the pen is equal to the difference between what John and Paul paid in total, and thus be able to find this 1-step algorithm. Lastly, the age problem (Table 1, right column) could also be solved using a 3-step algorithm ($17 - 8 = 9$; $8 - 2 = 6$; $9 + 6 = 15$) or a 1-step algorithm ($17 - 2 = 15$). Gamo et al. (2010) found that participants used the 1-step algorithm more often on such problems, even though all three problems were isomorphic and shared the exact same mathematical structure (see Fig. 1).

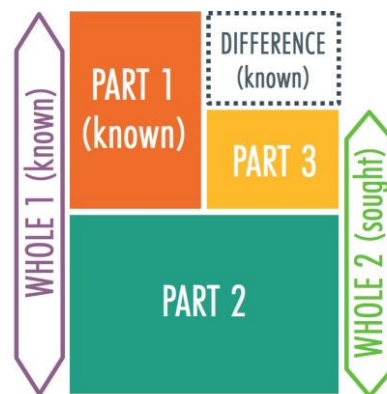


Fig. 1: The mathematical structure of the problems.

The authors argued that the reason behind this difference in participants' ability to use the shortest 1-step algorithm on all three problems was that the quantities used (family members, price, age) evoked different aspects of world knowledge that emphasized either the *cardinal* or the *ordinal* nature of the problems' values. The interpretation was that price or family problems tend to be encoded as sets of unordered, disconnected elements that can be grouped and whose values can be added with no ontological order (Gamo et al., 2010). Such an encoding would lead participants to calculate the individual value of each part making the whole, and then add them up, resulting in a 3-step algorithm. On the other hand, due to participants' experience of time represented along an oriented number line, the age values are seen as inherently ordered and can be represented as states (positions on an axis)

and transitions (intervals on a timeline). According to the authors, this type of encoding would make it easier to compare two sets sharing a common part: if Jean and Antoine started attending painting classes at the same age, and Antoine attended the class for 2 years less than Jean, then making the inference that Jean was 2 years younger than Antoine at the end of their respective classes seems especially natural.

Interestingly, one of the specificities of the problems used by Gamo et al. (2010) was that depending on whether participants constructed a cardinal or an ordinal encoding of the situation, some of the problems' values would become irrelevant information. Indeed, even though all 3 numerical values are needed to calculate the solution using the 3-step strategy, the value of Part 1 (see Fig. 1) becomes irrelevant for the 1-step strategy. The part played by the presence of such "conditionally irrelevant" information remains to be elucidated. Discriminating relevant from irrelevant information is known to be an arduous task (Carpenter et al., 1988; Kouba & Wearne, 2000; Low & Over, 1993). In fact, several studies have considered the influence of irrelevant information on participants' performance; according to Hembree's (1992) meta-analysis of 27 studies, lower performances are associated with the presence of "extraneous data" in the form of an irrelevant number in a problem. Besides, Cook and Rieser (2005) consider that irrelevant numerical information might be more difficult to handle than irrelevant contextual information, due to the similarity of the former to the relevant information. Similarly, Ng, Lee, and Khng (2017) also found that the use of irrelevant numbers in a problem may have a detrimental effect on performance, while contextual irrelevant information may not have such an effect. More recently, Vondrová (2020) brought converging evidence showing the negative influence of an irrelevant number on word problem difficulty. Considering how the presence of irrelevant numerical information in a problem statement was shown to impede performance, we decided to explore this issue in the last two experiments of our study, by using a modified version of the problems where the "conditionally irrelevant" value was removed, thus making it impossible to use the 3-step strategy.

Current study

Going beyond this preliminary work, our ambition for this paper is twofold. First, we aim at showing the critical influence that the distinction between cardinality and ordinality has on human understanding of situations involving numbers. We intend to demonstrate that, more than a question of learning how to enumerate collections, this is an ontological issue regarding the way numerosity is perceived. Second, we intend to build on this semantic distinction to investigate adults' representations in the course of mathematical reasoning. Namely, we aim at showing that the nature of the entities staged in word problems interferes with their classification, comparison,

encoding, recoding and solving, thus presenting a range of effects falling beyond the scope of the current models of arithmetic problem solving. We designed a series of 6 experiments scrutinizing how the cardinal versus ordinal aspects of numerical situations influence adults' reasoning in a variety of tasks.

Overview of the experiments' research questions

We intend to show that the selection of a specific type of quantity to create a problem statement has an influence on the encoding of the problem into either a cardinal or an ordinal representation, which in turn fosters the use of one of the two existing solving algorithms (see Fig. 2 for a graphical summary of this hypothesis). We designed six experiments in order to examine the importance that the cardinal versus ordinal distinction holds in the representation of numerical situations. They were conducted with adults to underline the strength and pervasiveness of the reported effects. Below, we propose a brief overview of the research questions investigated in each experiment. The specific hypotheses for each experiment will be further described in their respective sections.

First, we devised a free sorting experiment which aimed at showing the fundamental influence of the ordinal versus cardinal distinction on participants' spontaneous categorization of problems, while also assessing the validity of our choice of materials. Second, we used an analogy identification experiment to determine whether the hypothesized encoding of cardinal and ordinal problems would predict how likely participants are to perceive the isomorphism between superficially dissimilar problem statements. Third, we designed a direct comparison task to evaluate the robustness of the second experiment's findings regarding participants' inability to perceive the analogous structure shared by cardinal and ordinal problems. Fourth, we proposed a solving task with cardinal, ordinal, and "hybrid" problems (problems with "ordinalized" cardinal quantities) to demonstrate that participants' choice of a solving strategy depends on the semantics introduced in the problem. Fifth, we removed one of the problems' numerical values and used a solvability-assessment task to determine whether participants could solve problems whose unique solution was incompatible with their initial encoding of the problems. Sixth, we used a solution-validity-judgment task to assess whether participants' difficulty to find a problem's unique solution in the fifth experiment could be overcome by the direct presentation of a candidate solution to evaluate. Altogether, these experiments pursued several related research questions, as they intended to validate the distinction between cardinal and ordinal situations, to show its influence on the encoding of numerical situations even among adults well past their schooling years, and to evaluate whether its influence could be so robust as to interfere with participants' ability to use relatively basic arithmetic knowledge to solve 1-step arithmetic problems.

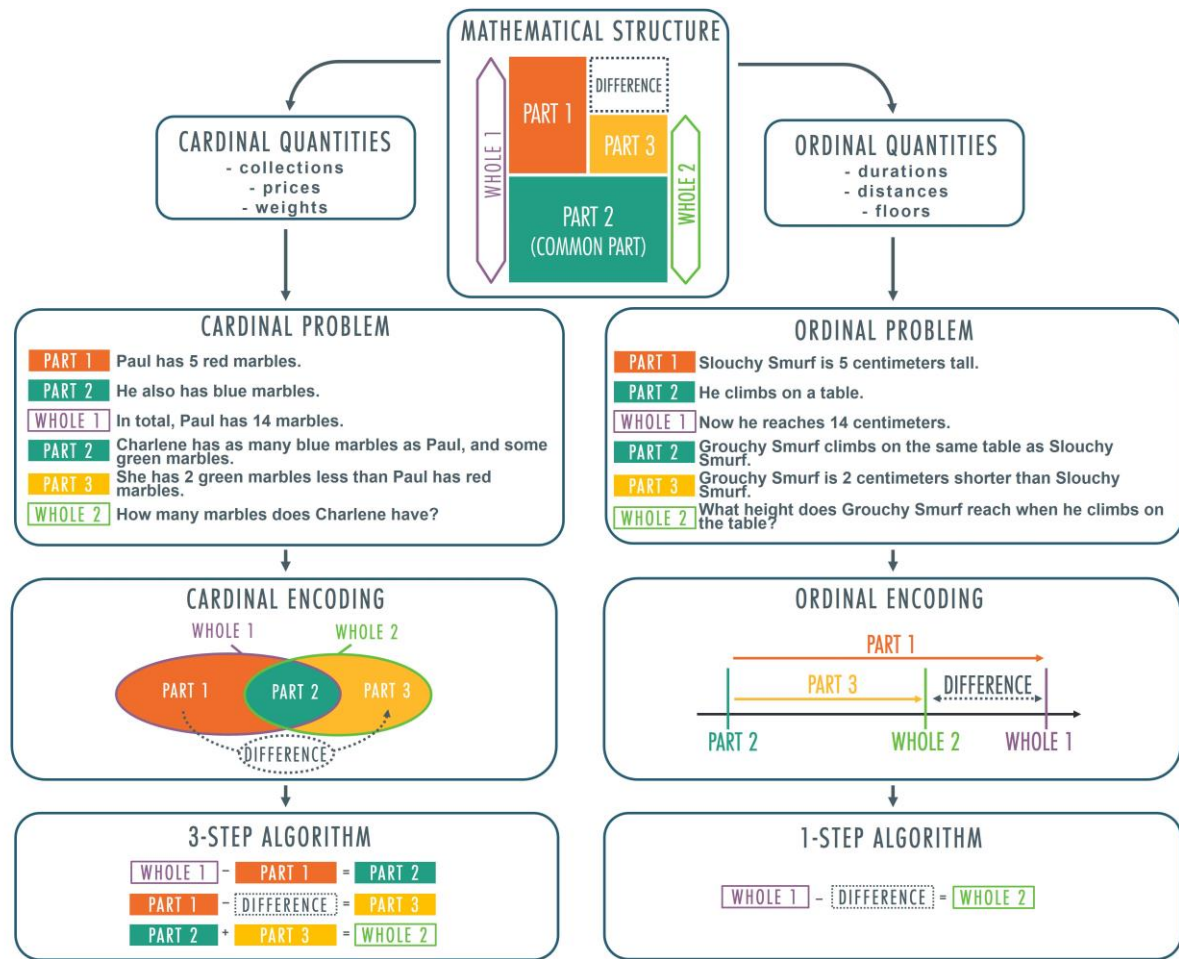


Fig. 2: Graphical summary of the content effects under scrutiny in this paper. Problems sharing the same mathematical structure are implemented with different quantities evoking different encodings, which in turn foster the use of one of the two solving algorithms. The problems all shared the same abstract mathematical structure, described in the “Mathematical structure” box, and the same pieces of information were provided in the same order across problems, as described in the “Cardinal problem” box and in the “Ordinal problem” box.

In an attempt to maximize the generality of our findings, we selected 3 types of quantities thought to evoke an ordinal encoding of the situation and 3 types of quantities that we assumed evoked a cardinal encoding. The rationale behind the selection of those specific quantities is presented below; the relevance of this classification will notably be assessed by our first experiment.

Overview of the selected ordinal quantities

Durations. In English, spatial metaphors expressing durations along a unidimensional space such as “a long time” are prevalent (Casasanto, 2008). The same is true for French, in which the word to designate an extensive period of time corresponds to the literal concatenation of the terms standing for “long” (“long”) and “time” (“temps”): “longtemps”. In most western cultures, the idea that time evolves along an axis or a timeline is deeply rooted in human understanding of the world (Bonato, Zorzi, & Umiltà, 2012; Boroditsky, 2011; Droit-Volet &

Coull, 2015; Weger & Pratt, 2008). Despite cultural differences regarding the direction given to the axis of time (e.g. Fuhrman & Boroditsky, 2010), the fact remains that time is usually conceived of as a unidirectional axis on which values are ontologically ordered.

Heights. Height is a spatial, unidimensional concept. Due to our living in a world where gravity constantly exerts its influence on physical objects, we quickly learn that objects fall in a straight line and we tend to think of height as being an oriented, vertical axis, with a bottom and a top (e.g. Hood, 1998; Hood, Santos, & Fieselman, 2000; Kim & Spelke, 1992). Entities' heights can easily be compared or stacked alongside this axis, and there is an ontological order in the values representing the heights of different entities placed atop each other.

Floors. Building floors are stacked on top of each other and they are ordered in a fashion that cannot be altered. If one is to imagine an elevator going from one floor to another, it immediately comes to mind that to go from the 1st floor to the 3rd floor, the elevator must pass the 2nd floor first (Clément & Richard, 1997).

Overview of the selected cardinal quantities

Collections. Collections refer to groups of disconnected elements that can be counted as parts of a set. Be it a collection of blue marbles in a bag, a set of iguanas in a terrarium or a group of pupils in a bus, collections of similar elements usually have no ontological order. Fittingly, most studies on the development of cardinality in children's early years resort to tasks consisting in counting collections of objects, such as the *Give-N* task (e.g. Condry & Spelke, 2008; Izard, Streri, & Spelke, 2014; Sarnecka & Carey, 2008).

Weights. In daily-life, weight is often seen as a property of some definite element, and as such it has no ontological order. For instance, when considering the weight of a stack of dictionaries, it does not matter which one is on top and which one is at the bottom of the stack: computing the total weight simply requires adding the weight of each individual book. Specific weights are assigned to specific entities, and although weight can vary over time, it is rarely encoded along an axis in daily life.

Prices. As with weight, price is usually considered the property of some unordered entities (Gamo et al., 2010). To calculate the total price of a series of items in a store, one usually adds the individual price of each item, in no specific order.

Problems wording

Of note, in addition to the reasons given above, our selection of these 6 different quantities to create the problems was also motivated by a previous study we conducted, using the same quantities, in which we asked participants to solve problems using these quantities and to make a drawing of the situation each problem depicted (Gros,

Thibaut, & Sander, 2017). In this study, we showed that the drawings of adults and 5th graders alike tend to include more ordinal features (e.g. axes, graduations, intervals) than cardinal features (e.g. sets, collections of items, value assignations) for problems using one of the 3 selected ordinal quantities, whereas they include more cardinal features than ordinal features for problems using one of the 3 selected cardinal quantities.

In the current study, we used those same 6 quantities to create a pool of 12 problems: 6 involved cardinal quantities (2 collection problems, 2 price problems, and 2 weight problems) and 6 featured ordinal quantities (2 duration problems, 2 height problem and 2 floors problems). Those problems were created specifically for the sake of this study, they were designed to meet three criteria: (i) they did not present arithmetic operations that a 4th grader would be unfamiliar with, (ii) they could be implemented either with cardinal quantities or with ordinal quantities, and (iii) they admitted two distinct solving strategies, each compatible with one of two encodings. Those problems were used as is in Experiments 1 to 3, while wording variations were introduced in Experiments 4 to 6. The problems were written in French (original materials for all experiments are available online at https://osf.io/kz6gh/?view_only=2f3fb6b910844e238ae58fef3c61168a. English translation is provided in Tables 2 and 3).

Table 2: Cardinal problems used in Experiment 1. The numerical values respected the following rule: $z < 4 < x < y < 15$.

Quantity used	Pb. ID	Problem statement
Weight	Pb. A	A bag of pears weighs x kilograms. It is weighed with a whole cheese. In total, the weighing scale indicates y kilograms. The same cheese is weighed with a milk pack. The milk pack weighs z kilograms less than the bag of pears. How much does the weighing scale indicate now?
	Pb. B	Tom takes a Russian dictionary weighing x kilograms. He also takes a Spanish dictionary. In total, he is carrying y kilograms of books. Lola takes Tom's Spanish dictionary and a German dictionary. The German dictionary weighs z kilograms less than the Russian dictionary. In total, how many kilograms is Lola carrying?
Price	Pb. C	In the first meal on the menu, there is a chocolate cake costing x euros. The meal also includes an omelet with mushrooms. In total, the first meal costs y euros. In the second meal on the menu, there is the same omelet with mushrooms, and an apple pie. The apple pie costs z euros less than the chocolate cake. How much does the second meal cost?
Price	Pb. D	In the stationery shop, Antoine wants to buy a x -euro ruler. He also wants a notebook. In total, that will cost him y euros. Julie wants to buy the same notebook as Antoine, and an eraser. The eraser costs z euros less than the ruler. How much will Julie have to pay?

Collection	Pb. E	Paul has x red marbles. He also has blue marbles. In total, Paul has y marbles. Charlene has as many blue marbles as Paul, and some green marbles. She has z green marbles less than Paul has red marbles. How many marbles does Charlene have?
Collection	Pb. F	Sarah owns x goldfish. Her other pets are all iguanas. In total, she owns y pets. Bobby is pet-sitting Sarah's iguanas during the holidays, he puts them with the turtles he owns. Bobby owns z turtles less than Sarah owns goldfish. How many pets are there at Bobby's?

Table 3: Ordinal problems used in Experiment 1. The numerical values respected the following rule: $z < 4 < x < y < 15$.

Quantity used	Pb. ID	Problem statement
Duration	Pb. G	The construction of the cathedral took x years. Before constructing it, the plans had to be made. The construction of the cathedral was completed in year y. The construction of the castle started at the same time as the construction of the cathedral. The construction of the castle took z years less than the construction of the cathedral. When was the construction of the castle completed?
Duration	Pb. H	Sophie travels for x hours. Her trip started during the day. Sophie arrives at y h. Fred leaves at the same time as Sophie. Fred's trip lasts for z hours less than Sophie's. What time is it when Fred arrives?
Height	Pb. I	Slouchy Smurf is x-centimeter tall. He climbs on a table. Now he reaches y centimeters. Grouchy Smurf climbs on the same table as Slouchy Smurf. Grouchy Smurf is z centimeters shorter than Slouchy Smurf. What height does Grouchy Smurf reach when he climbs on the table?
Height	Pb. J	Obelix's statue is x-meter tall. It is placed on a pedestal. Once on the pedestal, it reaches y meters. Asterix's statue is placed on the same pedestal as Obelix's. Asterix's statue is z meters shorter than Obelix's. What height does Asterix's statue reach when placed on the pedestal?
Floors	Pb. K	Naomi takes the elevator and goes up x floors. She left from the floor where her grandparents live. She arrives at the y th floor. Her brother Derek also takes the elevator from their grandparents' floor. He goes up z floors less than Naomi. What floor does Derek arrive at?
Floors	Pb. L	Karen takes the elevator and goes up x floors. She left from the floor where the gym is. She arrives at the y th floor. Yohan also takes the elevator from the floor where the gym is. He goes up z floors less than Karen. What floor does Yohan arrive at?

The problems all shared the same mathematical structure, they had the same number of sentences and their

numerical values were systematically mentioned in the same order. Cardinal and ordinal problems did not significantly differ in number of words; $t(10) = 0.98$, $p = .35$, independent t-test. The goal was to create cardinal and ordinal problems as comparable to each other as possible. The problems were isomorphic, and the numbers used were randomized across problems. The computation of the problems' solving algorithms only required basic mathematical knowledge (additions and subtractions of values below 15) so that the performance differences could not be explained by computational difficulty or lack of knowledge about the arithmetic operations involved.

Potential confounds and preliminary observations

The idea behind the current study was to elicit either a cardinal or an ordinal encoding of problems sharing the same mathematical structure. Instead of resorting to digit ordering or magnitude comparison tasks to evaluate cardinal and ordinal understanding separately, we intended to show the relevance of these two dimensions within complex situations worded using natural language. Thus, despite some unavoidable wording variations from one problem to another, the main difference between cardinal and ordinal problems was meant to be the use of cardinal versus ordinal quantities. We were especially careful to avoid potential confounds such as a “garden path” effect in which participants are led to make an erroneous deduction about a sentence or a problem's goal (Johnson, Moen, & Thompson, 1988; Tomasello & Herron, 1988). In our case, leading the participants “down the garden path” would have meant changing the problem statements' structure to lead participants to calculate irrelevant values in some versions of the problems but not in others. On the contrary, we made sure to present the same pieces of information regarding the mathematical structure of the problems in the same order and within the same sentences across problems. For instance, the second sentence of the problems always introduced the Part 2 of the problems, both in cardinal problems (e.g. “He also has blue marbles”) and in ordinal problems (e.g. “He climbs on a table”). Our intent was to make sure that if participants were to calculate the value of Part 2 on cardinal problems but not on ordinal problems, then it would not be possible to attribute this difference to the order in which the different parts and wholes are mentioned in the problems. Instead, we believe that such a difference would necessarily come from the semantic distinction we introduced between the cardinal and ordinal problem statements.

Similarly, we made sure that the difference between cardinal and ordinal problems would not boil down to a mere difference between the structural components (the schemata) of the problems. The schema theory (Kintsch, 1988; Kintsch & Greeno, 1985; Rumelhart, 1980; Schank, 1975; Schank & Abelson, 1977) proposes that our ability to identify the algorithmic solution of a word problem depends on the activation of the appropriate schema in long-term memory (e.g. a change schema, a combine schema, a compare schema). According to this theory,

understanding a word problem equates to triggering the relevant schema and implementing it with numerical information from the text (Verschaffel & De Corte, 1997). For instance, solving a problem such as “Pete has 3 apples. Ann also has some apples. Pete and Ann have 9 apples altogether. How many apples does Ann have?” would require possessing schematic knowledge about combine components and mapping the problem’s values onto it (De Corte & Verschaffel, 1985). With that in mind, and considering the prominence of the schema theory in the arithmetic word problem solving literature, we strove to create cardinal and ordinal problems that would not be confined to one specific schema, to avoid potentially problematic confounds. For instance, while the ordinal “number of floors” problem K (see Table 3) involves a change component (Naomi’s floor changes when she goes up the elevator), this is not necessarily the case for all ordinal problems. The ordinal duration problem G, for instance, does not involve any quantity changing. Instead, it features a combine component where the duration of the cathedral’s construction is combined to the time needed to make the plans. Similarly, the cardinal collection problem E (see Table 2) involves combine components (with sets of marbles being put together), whereas the cardinal weight problem A also involves a value changing over time (the weight indicated by the weighing scale). Additionally, all the problems featured a compare component in the problem’s fifth sentence. In other words, while the schema theory is an interesting framework to analyze some of the difficulties one can meet while attempting to solve one-step additive word problems, we believe that the cardinal versus ordinal distinction we introduced between these problems is not simply a matter of which schemata they could be interpreted through, but rather a question of how one will encode a given problem involving ordinal or cardinal quantities.

Similarly, we believe that the distinction between ordinal quantities and cardinal quantities is not a matter of how continuous they are, but rather of *how likely* they are to emphasize, respectively, the order between the numerical values or their individual magnitudes. Although the difference between discrete and continuous quantities is relevant in arithmetic word problems and deserving of attention (see Bagnoud, Burra, Castel, Oakhill, & Thevenot, 2018), we believe it is a separate issue. In fact, quantities qualifying as “discrete” were featured both in cardinal problems (“collections” problems) and in ordinal problems (“number of floors” problems) and quantities usually qualifying as “continuous” were used both in cardinal problems (“prices” and “weights” problems) and in ordinal problems (“durations” and “heights” problems). Thus, no matter the relative continuity or discreteness of the numerical dimension used, the main factor we manipulated between cardinal and ordinal problems remained the cardinal-ordinal distinction.

Experiment 1

The first experiment, a problem sorting task, was designed to assess whether the assumed difference between cardinal and ordinal quantities would lead to different encodings of the problems, reflected by categorization patterns consistent with this distinction. In other words, it intended to provide evidence for the existence of a fundamental distinction between cardinal and ordinal encodings derived from problems sharing an identical deep structure but differing in the quantities used in their problem statements. Indeed, we expected participants' categories to reveal the hypothesized difference between cardinal and ordinal problem statements. The experimental design was modeled on the work from Chi, Feltovich, and Glaser (1981), who performed a series of sorting experiments with experts and novices. They showed that, when asked to sort physics problems, novices put together problems sharing similar surface features (e.g. problems featuring pulleys), whereas participants with higher proficiency in physics favored the use of abstract physics principles (e.g. problems that can be solved using Newton's second law) to sort the problems.

Our prediction regarding such problem statements was that the adults' understanding of the problems would neither be strictly limited to literal features, as was the case for Chi et al.'s (1981) lay participants, nor would it be guided by the mathematical deep structure of the problems. Rather, we hypothesized that participants would sort these problems sharing the same mathematical structure depending on whether they encode them as emphasizing ordinality versus cardinality. We predicted that participants would tend to group together problems evoking a cardinal encoding (collection, price, and weight problems) and that they would group together problems evoking an ordinal encoding (duration, height, and number of floors problems), regardless of the problems' other features. In other words, despite collection problems being as dissimilar from price problems than from duration problems with regards to their surface features, we expected participants to group collection problems with price problems more often than they would with duration problems, due to them constructing a cardinal encoding for collection and price problems and an ordinal encoding for duration problems."

Method

Participants. Participants were recruited on a voluntary basis in the university library. Data collection took place over the course of one week, at the end of which a total of 85 participants had agreed to participate in the experiment (54 women and 31 men, $M = 24.31$ years, $SD = 8.33$). All participants spoke French fluently. None had previously participated in any similar experiment.

Materials and procedure.

In this experiment, each of the 12 word problems was printed on a separate card. As in Chi et al.'s (1981) experiment, the task was to sort the problems into groups based on similarities of solving strategies. The following instructions were given to the participants:

Here are 12 arithmetic word problems. Please read and study carefully each of them. Your task is to sort the problems into groups based on similarities of solution. You can make as many groups of problems as you deem necessary. This is not a speed test: take your time to read and understand each of these problems. *Translated from French.*

We chose to use instructions similar to those given by Chi et al. (1981), in that we did not allow participants to use a pen and paper so they were less tempted to actually solve the problems before sorting them, nor did we specify a set number of problem categories to be made. We hypothesized that the cardinal versus ordinal distinction would be so influential that it would tend to appear even if participants were not explicitly instructed to sort the problems into only two categories. Participants all completed the task in less than an hour.

Results

The categories created by each participant were coded with a co-occurrence matrix describing how many times two problems were sorted together within the same category. A proximity matrix was then elaborated based on the co-occurrence matrix, describing the average perceived proximity between each problem (see Fig. 3). This matrix specifies which problems were sorted together most frequently; the higher the value between two problems, the higher the proportion of participants who considered these two problems similar. We tested the hypothesis that problems evoking a similar encoding (either ordinal-ordinal pairs or cardinal-cardinal pairs) would be categorized together more frequently than problems whose encoding is dissimilar (pairs made of one cardinal and one ordinal problem). Results showed that cardinal-cardinal pairs were given a higher proximity score on average ($M = 0.90$, $SD = 0.03$) than cardinal-ordinal pairs ($M = 0.75$, $SD = 0.02$); $t(49) = 21.31$, $p < .001$, independent t-test. Similarly, ordinal-ordinal pairs were attributed a higher proximity score on average ($M = 0.91$, $SD = 0.04$) than cardinal-ordinal pairs; $t(49) = 20.60$, $p < .001$, independent t-test. In fact, the perceived proximity between any pair of problems whose hypothesized encoding is similar (ordinal-ordinal or cardinal-cardinal) was systematically higher than the perceived similarity between any pair of problems whose hypothesized encoding is dissimilar (cardinal-ordinal).

		Weight problems		Cardinal problems		No. of elements problems		Duration problems		Ordinal problems		No. of floors problems	
		pb. A	pb. B	pb. C	pb. D	pb. E	pb. F	pb. G	pb. H	pb. I	pb. J	pb. K	pb. L
Cardinal problems	Weight problems	pb. A	-	0.95	-	0.90	0.89	0.88	0.86	0.79	0.78	0.76	0.75
		pb. B	0.95	-	0.87	0.94	0.90	0.89	0.75	0.74	0.75	0.72	0.72
	Price problems	pb. C	0.90	0.87	-	0.91	0.90	0.86	0.76	0.75	0.77	0.77	0.78
		pb. D	0.89	0.94	0.91	-	0.87	0.89	0.75	0.74	0.76	0.74	0.74
	Collection problems	pb. E	0.88	0.90	0.90	0.87	-	0.95	0.74	0.71	0.74	0.74	0.75
		pb. F	0.86	0.89	0.86	0.89	0.95	-	0.75	0.73	0.76	0.72	0.74
Ordinal problems	Duration problems	pb. G	0.79	0.75	0.76	0.75	0.74	0.75	-	0.95	0.86	0.87	0.89
		pb. H	0.76	0.74	0.75	0.76	0.71	0.73	0.95	-	0.87	0.88	0.91
	Height problems	pb. I	0.78	0.75	0.77	0.76	0.74	0.76	0.86	0.87	-	0.95	0.92
		pb. J	0.78	0.75	0.77	0.74	0.74	0.72	0.87	0.88	0.95	-	0.91
	Floors problems	pb. K	0.75	0.72	0.77	0.74	0.74	0.74	0.89	0.91	0.92	0.91	1.00
		pb. L	0.76	0.72	0.78	0.74	0.75	0.74	0.89	0.91	0.93	0.91	-

Fig. 3: Co-occurrence proximity matrix between problems. A higher number (and darker hue) indicates a higher co-occurrence frequency.

To better understand participants' classification patterns, we undertook a hierarchical cluster analysis displaying the global taxonomy of the collected categories, using the R package *pvclust* (Suzuki & Shimodaira, 2006). Fig. 4 details the clusters appearing in the dataset. For each cluster, the Approximately Unbiased (AU) p-value can be interpreted as follows: if $AU > .95$, the hypothesis that “the cluster does not exist” can be rejected at the significance level of .05. This suggests that these clusters do not reflect clustering noise and may be observed in a stable manner.

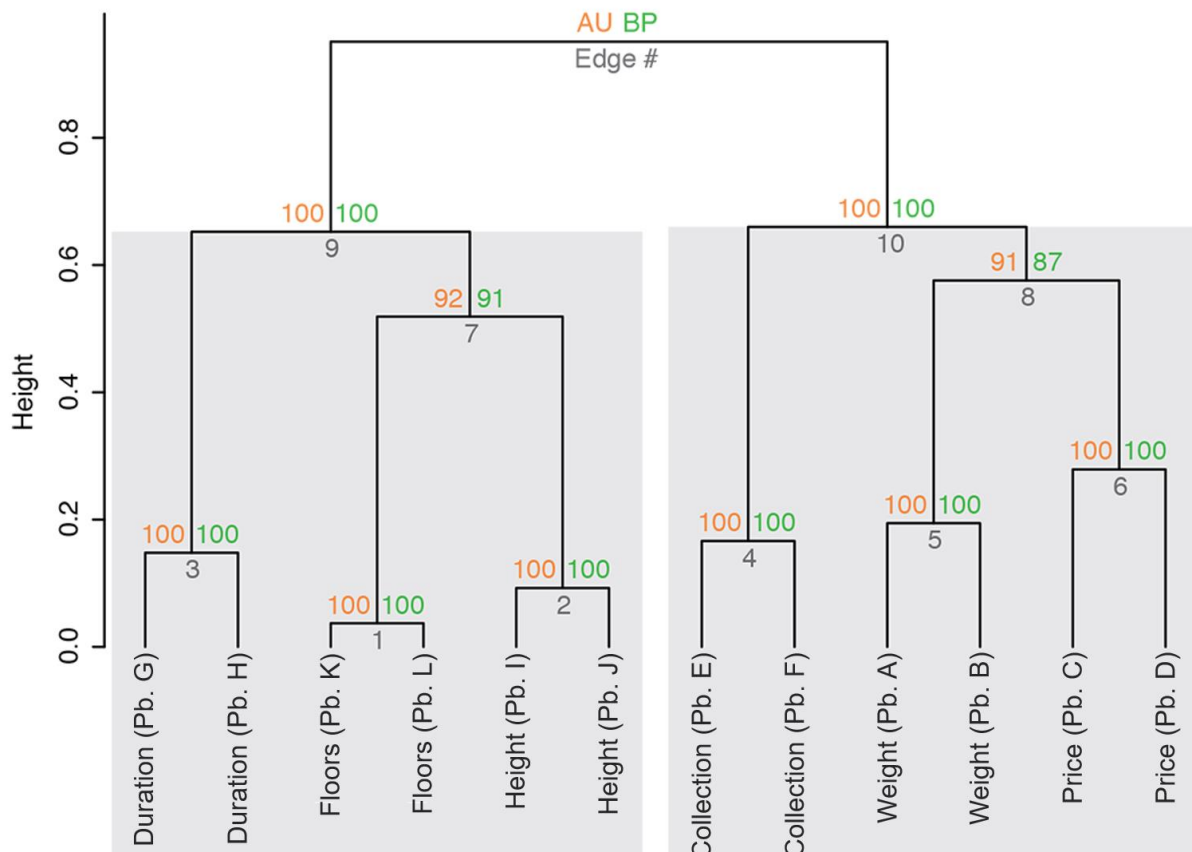


Fig. 4: Cluster dendrogram of the problems. Values are Approximately Unbiased (AU) p-values (orange, left) computed by multiscale bootstrap resampling, Bootstrap Probability (BP) values (green, right) computed by normal bootstrap resampling, and cluster labels (grey, bottom). Clusters with $AU \geq 95$ are highlighted by the grey rectangles and are considered to be strongly supported by the data. Number of bootstrap samples = 10,000.

All the problems sharing the same quantities (e.g. the two weight problems, the two duration problems, etc.) were grouped in the same lower clusters. This was expected since they both evoked the same encoding (either cardinal or ordinal) and shared a certain degree of surface similarity as well (two problems involving objects being weighed, for example, will both use words such as “weight”, “scale” or “weighs”, increasing the number of identical surface features). Yet, the two higher clusters displaying a significant AU p-value respectively regrouped all the ordinal problems and all the cardinal problems, regardless of their more specific surface features. This result supports the hypothesis that the cardinal versus ordinal distinction guided participants’ sorting patterns. Thus, as hypothesized, some aspects of the problem statements that pertain neither to the most specific surface features nor to the deep structure of the problems influenced the classification. The Bootstrap probability (BP) values, although presumably more biased (Suzuki & Shimodaira, 2006), confirmed this pattern in the data.

Discussion

This experiment sought to establish that participants’ categories were guided by the semantic distinction we introduced between cardinal and ordinal quantities. As predicted, the results showed that participants’ categories were not just based on the most superficial similarities between problems. Had it been the case, duration problems would not have been considered as closer from number of floors problems than from collection problems, for instance. Thus, even if participants did pick up on certain literal features of the problems, the cardinal/ordinal distinction remained significant even when other classification patterns (e.g. pairing problems based on their specific quantities, and nothing else) could have overshadowed them. Instead, participants perceived the similarity between problems fostering a similar encoding: problems evoking sets of unordered elements were grouped together, and so were problems with elements that could be ordered along an oriented axis. Of note, none of the participants proposed less than 2 categories, although this may be due to participants’ expectations regarding the minimum number of categories that could be made, rather than to them failing to notice that all problems shared the same mathematical deep structure (what Bassok, 2001, calls the objective mathematical structure). Overall, the results supported our hypothesis that weight, price, and collection problems are encoded differently than height, duration, and floor problems, thus substantiating our selection of those two sets of quantities. Here, the cardinal versus ordinal distinction significantly guided the sorting task, thus corroborating the fundamental role of this distinction.

Experiment 2

To evaluate how cardinality and ordinality interact with one another when different situations are being compared,

and to gather converging evidence regarding the influence that this distinction holds on adults' apprehension of numerical situations, our second experiment focused on participants' interpretation of different problems as analogous. Participants were presented with an unsolved word problem and asked to determine whether a series of target word problems could be solved analogously. We tested the hypothesis that participants can more easily perceive an analogy between two isomorphic problems if they feature quantities evoking a similar encoding (two problems emphasizing the cardinal nature of numbers, or two problems emphasizing the ordinal nature of numbers) than if they do not.

Methods

Participants. Participants were recruited as part of an undergraduate class at the University of Paris VIII. A total of 191 students (116 women and 75 men, $M = 27.3$ years, $SD = 11.9$) agreed to take part in the experiment on a voluntary basis. All participants spoke French fluently. None had previously participated in any similar experiment.

Materials and procedure. In this experiment, we used the same problems as those created for Experiment 1, with the addition of one cardinal problem and one ordinal problem (see complete materials online). Each participant was given a 4-page booklet. On the first page, the instructions read:

Below is an arithmetic word problem. Please read it and then study carefully the problems presented on the next page. Note, for each of them, if they can be solved using the same solving principle as the problem presented below. This is not a speed test: take your time to read and understand each of these problems.

Translated from French.

On the same page, a problem statement was printed ("problem A"). This problem was either a cardinal problem (a collection problem) or an ordinal one (a duration problem). On the following page, 6 target problems were presented: 3 ordinal problems (duration, height, floors), and 3 cardinal problems (collection, price, weight). Next to each problem, the participants had to circle their answer "yes" or "no" to the question "can this problem be solved similarly to problem A?". The next two pages had the same setup with different problem statements, where "problem A" was replaced by "problem B", that was either a collection problem (if "problem A" had been a duration problem) or a duration problem (if "problem A" had been a collection problem), and six new target problems.

Like in Experiment 1, we based our instructions on Chi et al.'s (1981) study and asked participants to compare problems based on their solving principle, while not giving them any other incentive to solve the problems. Additionally, participants were not given additional paper to write on, to discourage them from engaging in the

resolution of every problem before making their choice. The order of problems A and B was randomized between booklets, as was the order in which the target problems were presented. Participants all completed the task in less than an hour.

We predicted that the participants' answers would depend on the similarity between the type of representations fostered by the source and target problems. Namely, participants should perceive the analogy between two cardinal problems or between two ordinal problems more easily than between a cardinal and an ordinal problem. This should translate into an interaction between the cardinal versus ordinal nature of the source problem and that of the target problems.

Results

We computed the rate of detection of similarity between the source and the target problems, depending on the nature of the quantities involved (see Fig. 5). A two-way repeated measures ANOVA was conducted on the rate of perceived similarity with nature of the quantity in the source problem (ordinal or cardinal) and nature of the quantity in the target problem (ordinal or cardinal) as within factors. As expected, there was no main effect of the cardinal or ordinal semantics attached to the source problem ($F(1,190) = 1.07$, $p = .30$, $\eta_p^2 = .01$) nor of the semantics attached to the target problem ($F(1,190) = 0.10$, $p = .08$, $\eta_p^2 < .01$). There was, however, a significant interaction effect between the two factors, indicating that cardinal target problems were more likely to be selected by the participants when a cardinal source problem was presented, and that ordinal target problems were more likely to be chosen when the source problem was ordinal as well ($F(1, 190) = 72.20$, $p < .001$, $\eta_p^2 = .28$). In addition to the interaction, we performed 2-by-2 comparisons of the rate of perceived similarity between cardinal and ordinal target problems, depending on the semantics imbued in the source problem. Results showed that cardinal target problems were judged analogous to cardinal source problems significantly more often than ordinal target problems (82.4% for cardinal-to-cardinal analogy; 60.9% for cardinal-to-ordinal analogy; $t(190) = 6.50$, $p < .001$, $\eta_p^2 = .18$, paired t-test). Similarly, ordinal target problems were judged analogous to ordinal source problems significantly more often than cardinal target problems (84.8% for ordinal-to-ordinal analogy; 62.1% for ordinal-to-cardinal analogy; $t(190) = 7.25$, $p < .001$, $\eta_p^2 = .22$, paired t-test). In other words, participants identified collection source problems as analogous to other cardinal problems more frequently than they did to ordinal target problems. Reciprocally, the duration source problems were more frequently perceived as analogous to other ordinal problems than to cardinal problems.

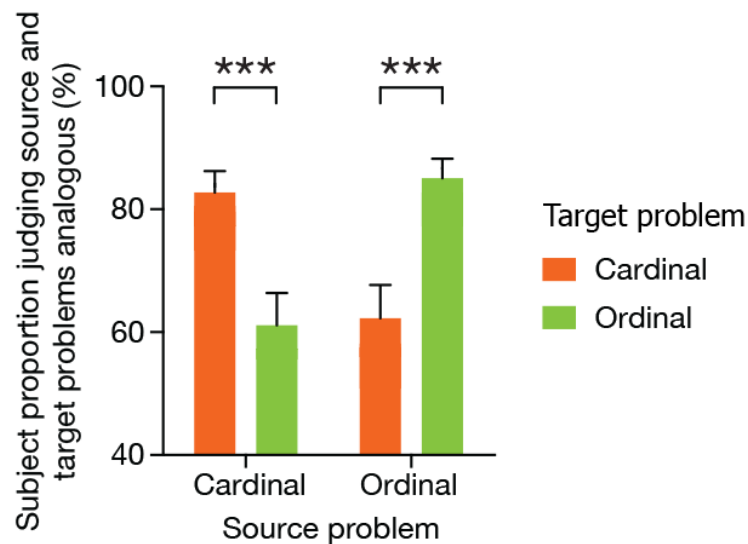


Fig. 5: Rate of perception of the analogy between the source and target problems, depending on the nature of their quantities. Vertical bars denote .95 confidence intervals. *** $p < .001$, paired t -test.

Overall, the analogy rate was higher when the source and target problems both evoked a similar encoding (either two cardinal encodings or two ordinal encodings) than when they evoked dissimilar encodings (a cardinal encoding and an ordinal encoding). Our hypothesis regarding the influence of the encoding on the detection of similarity of solving pattern was thus supported: participants' decisions were significantly influenced by the semantics attached to the quantities used in the problems.

Discussion

With this experiment, we showed that the encoding difference between cardinal and ordinal problems influences the perception of analogies between isomorphic problems. Participants had significantly more difficulties identifying that two problems shared the same solving principle when these problems featured elements evoking different aspects of their knowledge about the world. This corroborates our claim that general abstract semantic properties, such as cardinality or ordinality, play a crucial role in the encoding of mathematical word problems. Depending on the semantics evoked by the problems, participants encode different representations. The representations in turn limit or foster participants' identification of the analogous relations between the problems. Our results support the idea that participants encode a representation whose nature depends on the world knowledge evoked by the problem statement.

While the interaction between the cardinal versus ordinal nature of the target and source problems indicates that participants' encoding of the problems was significantly influenced by the quantities they featured, it might be that participants could have overcome this inability to perceive the similarity between cardinal and ordinal problems if they had tried to directly map the structures of the problems to one another. In other words,

suppose participants were encouraged to compare one specific target problem with the source problem instead of simply being asked to select the analogous problems among a series of potential candidates. Would they go beyond their initial encoding of the situation and identify the isomorphism between the problems? We designed a third experiment to answer this question.

Experiment 3

Experiment 3 aimed at replicating Experiment 2's findings and assessing their robustness by directly presenting pairs of problems to the participants, instead of asking them to identify among a list of target problems the ones that shared a solution principle with the source problem. We hypothesized that the effect observed in Experiment 2 could be replicated in a one-to-one comparison setting, in which participants are encouraged to directly contrast two problem statements. We assumed that when the source and the target problems featured the same type of quantity – cardinal or ordinal – participants would acknowledge the solution equivalence more often than when problems featured different types of quantity.

Methods

Participants. This experiment was conducted online, on the survey platform Qualtrics. Survey link was sent through social networks and by email. We decided to keep the survey open for one week. After one week, a total of 147 adults participated in the experiment (60 women and 87 men, $M = 30.0$ years, $SD = 11.5$). We estimated that the sample collected was close enough to our target sample size. All participants spoke French fluently. None had previously participated in any similar experiment. The results of five participants were removed since they failed to provide an answer to one or more questions in the experiment. The analyses bear on the 142 remaining participants (59 women and 83 men, $M = 29.1$ years, $SD = 10.2$ years).

Materials and procedure. This experiment was conducted on the Qualtrics platform for online experiments. The source and target problems were the same as those used in Experiment 2. On the first page, the instructions read:

Below is an arithmetic word problem. Please read it carefully. On the next pages, you will be presented with a series of arithmetic problems. Indicate, for each new problem, whether it can be solved with the same solving principle as the problem presented below. This is not a speed test: take your time to read and understand each of these problems. *Translated from French.*

A source problem was then presented, evoking either a cardinal or an ordinal encoding. The following 6 pages repeated the source problem, and then presented a new problem below. Each time, the following question was

displayed: “Can these two problems be solved using a similar solution principle?”. After 6 target problems had been introduced, a new source problem was presented (a cardinal problem if the first target problem was ordinal, an ordinal problem if the first target problem was cardinal), with the same instructions as before, successively followed by 6 new target problems, one at a time. The source and target problems were the same as those used in Experiment 2. Like in Experiment 2, there was no explicit incentive to actually solve the problems, since they had been told beforehand to refrain from using outside materials (pocket calculator, pen and paper) while participating in the online experiment. Participants all completed the task in less than an hour.

Results

For each type of source problem, we analyzed the percentage of participants answering that the source and target problems might be solved following a similar solution principle (see Fig. 6). A two-way repeated measures ANOVA was conducted on participants’ rate of identified similarity, with nature of the quantity in the source problem (cardinal or ordinal) and nature of the quantity in the target problem (cardinal or ordinal) as within factors. As in Experiment 2, there was no main effect of the cardinal or ordinal semantics attached to the source problem ($F(1,140) = 2.14, p = .15, \eta_p^2 = .02$) nor of the semantics attached to the target problem ($F(1,140) = 0.41, p = .53, \eta_p^2 < .01$). However, we replicated the interaction observed in Experiment 2 between the semantic nature of the source problems and that of the target problems ($F(1, 140) = 73.39, p < .001, \eta_p^2 = .34$).

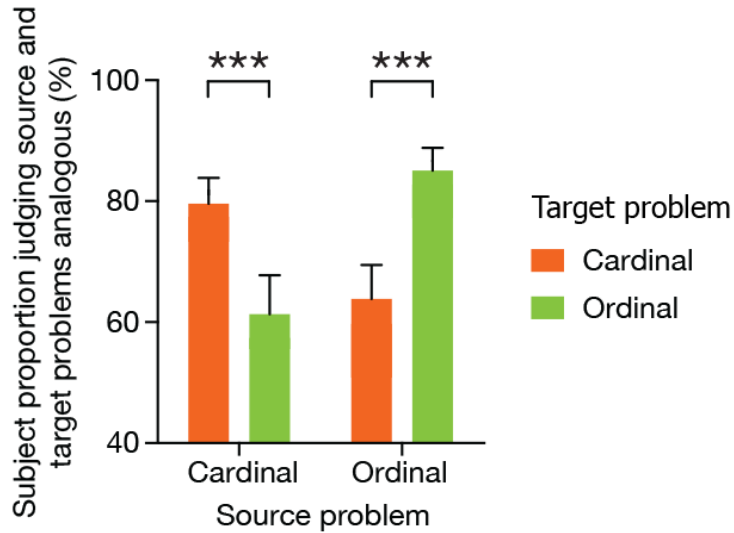


Fig. 6: Rate of perception of the analogy between the source and target problems, depending on the nature of their quantities. Vertical bars denote .95 confidence intervals. *** $p < .001$, paired t-test.

In addition to the interaction, 2-by-2 analyses revealed that cardinal target problems were judged analogous to cardinal source problems significantly more often than ordinal target problems (80.9% for cardinal-to-cardinal analogy; 62.6% for cardinal-to-ordinal analogy; $t(140) = 5.51, p < .001, \eta_p^2 = .18$, paired t-test). Similarly, ordinal

target problems were judged analogous to ordinal source problems significantly more often than cardinal target problems (87.7% for ordinal-to-ordinal analogy; 65.7% for ordinal-to-cardinal analogy; $t(140) = 6.91$, $p < .001$, $\eta_p^2 = .25$, paired t-test). In other words, despite participants being encouraged to directly compare one source problem with one target problem, the problems remained classified as analogous based on the semantics they carried.

Discussion

In this experiment, it was showed that the effect of the distinction between cardinal and ordinal quantities on the encoding of the situations described in the problems was not altered by the direct presentation of two problems side by side. As predicted, participants had significantly more difficulties identifying that ordinal-cardinal pairs of problems could be solved in the same way, even though the simultaneous presentation of both problems should have facilitated their mapping. Interestingly, the results of Experiment 2 were replicated in this new experimental setting. Our hypothesis regarding the influence of the cardinal versus ordinal distinction on the detection of similarity of solving pattern was thus supported by both experiments. One question that follows from these two experiments regards the role of these robust encoding mechanisms in the choice of a solving strategy. To what extent do the constructed representations dictate participants' solving strategies? By using problems designed to evoke cardinal, ordinal, or hybrid encodings, the next experiment was designed to evaluate how these encoding differences may influence participants' choice of a solving strategy in a more traditional solving task.

Experiment 4

The goal of this fourth experiment was twofold. First, we intended to demonstrate that participants' ability to use a specific solving strategy directly depended on the nature of their semantic encoding of the problem. Second, we aimed at showing that by changing the semantics imbued in a problem statement, we could significantly alter its encoding on the cardinal versus ordinal dimension. To this end, we introduced a new type of problems, in addition to the cardinal and ordinal problems used in the previous experiments. We called those new problems "hybrid", as they were meant to elicit an ordinal encoding of cardinal quantities by adding an ordinal dimension to the situation; a change occurring over time. Indeed, those hybrid problems involved the same cardinal quantities as the cardinal problems (prices, weights, collections) but their problem statements featured a scenario fostering an ordinal encoding by describing *how the cardinal quantities changed over time*. For instance, hybrid weight problems were created by describing the weight of a baby growing over time in order to favor an ordinal representation of the weight. We predicted that introducing those characteristics without changing the quantities

themselves nor the problems' structure would influence the encoding of the problems and the strategies subsequently implemented.

Since we instructed participants to solve the problems using as few operations as possible, our main hypothesis regarded the rate of use of the 1-step strategy. We predicted that problems involving ordinal quantities would lead to a greater use of the 1-step strategy than problems involving cardinal quantities, due to the ordinal encoding making it easier to perceive the validity of this solution. Second, we hypothesized that hybrid problems would lead to a significantly higher rate of 1-step strategy than cardinal problems, due to the ordinal semantics attached to the problem statements. Additionally, we aimed at assessing whether hybrid problems would be solved by the 1-step strategy as often as ordinal problems, or not.

Method

Participants. Participants were recruited on a voluntary basis among the undergraduate population of three universities. A total of 181 students participated in this experiment after giving informed consent (123 women and 58 men, $M = 23.35$ years, $SD = 7.82$). All participants spoke French fluently. None had previously participated in any similar experiment.

Materials and procedure. A pool of 18 word problems was used for this experiment: the same 12 problems (6 cardinal and 6 ordinal) as in the first three experiments, and 6 new hybrid problems (see Table 4 for a selection of hybrid problems). Each participant saw 9 problems in total: 3 of each category.

Table 4: Example of hybrid problems. The numerical values respected the following rule: $z < 4 < x < y < 15$.

Quantity used	Hybrid problem statement
Weight	"During his first year, David gained x kilograms. At birth, he already weighed a certain weight. After a year, David weighs y kilograms. At birth, David and Lara had the same weight. During her first year, Lara gained z kilograms less than David did. How much does Lara weigh after one year?"
Price	"For Christmas, Felix got x euros. He already had some money. Now Felix has y euros in total. Before Christmas, Zoe had as much money as Felix. For Christmas, Zoe got z euros less than Felix did. How much money does Zoe have now?"
Collection	"During the afternoon, Patricia catches x fish. She puts those fish in her basket, with the other fish she caught during the morning. By the end of the day, Patricia has y fish in her basket. During the morning, Arthur caught as many fish as Patricia did. During the afternoon, Arthur catches z fish less than Patricia does. In total, how many fish did Arthur catch today?"

The participants all received 10-page booklets with instructions printed on the first page. The instructions read:

You will find an arithmetic problem on each page of this booklet. Your task is to solve the problems using as few operations as possible. You can use the ‘draft’ area, but please copy in the ‘response’ area all the operations that you used to come up with the solution. This is not a speed test: take your time to read and understand each of these problems. Remember that the goal is to solve the problems using as few operations as possible. For every problem, we ask you to write down every operation(s) that you used to come up with the solution, even the simplest one that you can calculate mentally. For instance, the computation “ $15 - 6 - 2 = 7$ ”, should not be written as a unique operation, but broken down as “ $15 - 6 = 9$ ” and “ $9 - 2 = 7$ ”, which then counts for two operations. *Translated from French.*

Problem order was randomized across booklets. Each page in the booklet was divided in three parts: the problem statement, the “draft” area and the “response” area. Participants all completed the task in less than an hour.

Scoring. A problem was considered as correctly solved when the obtained result came with the appropriate calculations. The strategies leading to success were categorized either as correct 1-step strategy or as correct 3-step strategy. When the written operations were correct and the written solution was within ± 1 of the correct result, this was deemed a calculation error and problems were still considered as correctly solved. Other answers were considered as false.

Results

The percentage of correct solving using both strategies was calculated. Fig. 6 details the rate of use of the 1-step strategy for each problem category. A one-way repeated measures ANOVA was conducted on the rate of use of the 1-step strategy between cardinal, hybrid and ordinal problems. Results indicated that there was a main effect of problem category (cardinal/hybrid/ordinal) ($F(2,360) = 52.13, p < .001, \eta_p^2 = .22$). Three paired sample t-tests were used to compare the different conditions, with Bonferroni adjustment for multiple comparisons. As predicted, cardinal problems led to a significantly lower rate of 1-step strategy ($M = 0.28, SD = 0.39$) than ordinal problems ($M = 0.47, SD = 0.39$); ($t(180) = 9.34, p < .001, d = 0.48$), which supported the hypothesis that the use of cardinal versus ordinal quantities significantly influenced the encoding of the problem statements, and subsequently shaped the solving strategies used by the participants. Moreover, as hypothesized, the 1-step strategy was more frequently used on hybrid problems ($M = 0.38, SD = 0.40$) than on cardinal problems; ($t(180) = 5.91, p < .001, d = 0.24$). This result showed that, by giving specific semantic properties to a problem statement, it was possible to influence the encoding and manipulate which solving strategy participants would use. A semantically cardinal quantity presented in an ordinal context could thus lead more often to the encoding of an ordinal

representation than a cardinal quantity presented in a context that does not present ordinal features.

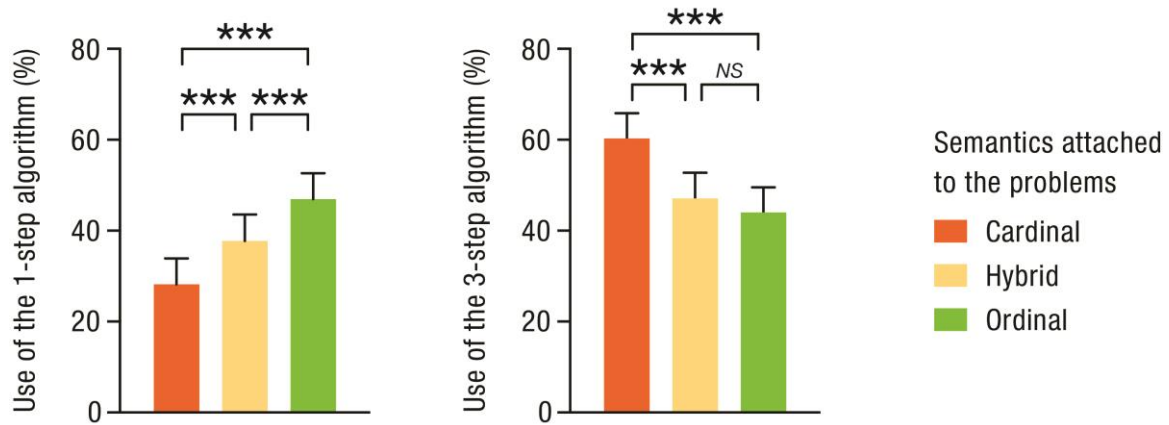


Fig. 6: Rate of use of the 1-step strategy and 3-step strategy, depending on the semantics attached to the problems. Vertical bars denote .95 confidence intervals. *** $p < .001$ (paired t-test).

Interestingly, the comparison between hybrid problems and problems with ordinal quantities revealed that the rate of use of the 1-step strategy was still higher on ordinal problems ($M = 0.47$, $SD = 0.39$); $t(180) = 4.93$, $p < .001$, $d = 0.23$. This seems to indicate that, while the manipulation that was performed to “ordinalize” cardinal quantities had a significant impact when compared to standard cardinal problems, an “ordinalized” quantity remained less ordinal than a typical ordinal one.

Regarding the rate of use of the 3-step strategy, we performed a one-way repeated measures ANOVA to evaluate if it differed between cardinal, hybrid and ordinal problems. Results indicated the presence of a main effect of problem category (cardinal/hybrid/ordinal) on the use of this strategy as well ($F(2,360) = 27.80$, $p < .001$, $\eta_p^2 = .13$). We used paired sample t-tests to perform pairwise comparisons between the three conditions, with Bonferroni adjustment for multiple comparisons. Results showed that participants resorted to the 3-step strategy more often on cardinal problems ($M = 0.60$ $SD = 0.39$) than on ordinal problems ($M = 0.44$ $SD = 0.38$); $t(180) = 6.79$, $p < .001$. Interestingly, they also used the 3-step strategy more often on cardinal than on hybrid problems ($M = 0.47$ $SD = 0.38$); $t(180) = 5.80$, $p < .001$. This is in line with the idea that introducing cardinal quantities in an ordinal context could help “ordinalize” participants’ representation. There was, however, no significant difference between the rate of use of the 3-step strategy on hybrid and on ordinal problems; $t(180) = 1.38$, $p = .17$.

Discussion

In this experiment, the analysis of the solving strategies provided cues on how the encoding of the problems influences participants’ solving strategies. Participants’ use of the shortest strategy was dependent on the type of quantities involved in the problem statement. Despite participants being explicitly instructed to use as few operations as possible to solve the problems, they mostly used the 3-step strategy to solve the cardinal problems.

This experiment further supported the claim that the ordinal versus cardinal dimension was the main factor constraining strategy choice. Indeed, a change in strategy choice followed the “ordinalization” of cardinal quantities, thus showing that the ability to use the 1-step strategy was directly dependent on how much the problem statement emphasized the ordinal nature of its numerical values.

So far, we have studied how arithmetic word problems are initially encoded and how the interpreted representations in turn influence the solving strategies used. However, as mentioned in the introduction, the initial encoding of a problem does not always provide a solution to the solver. We then investigated whether participants can overcome their initial representation of the problems when the one they first encoded does not lead to a solution. That is, we created the conditions to explore the difference between situations in which one has to construct a new representation of the situation presented, and those in which there is no need for a new representation to be constructed. Such a recoding process would imply to disregard the cardinal semantics evoked by specific quantities, and to construct a new encoding of the situation regardless of their influence.

Experiment 5

In this fifth experiment, we designed situations meant to have participants construct a representation leading to a dead end. In other words, we created problems for which the initial encoding would not provide a successful solving strategy, thus hindering the solving process. The problems could nonetheless be solved if participants constructed a different encoding of the situation. We tested the participants’ proficiency to use the 1-step solving strategy on problems that would spontaneously elicit the 3-step strategy. For that purpose, cardinal problems that only featured two numerical values were introduced, making the 3-step solving strategy impossible to use. By contrast, the 1-step strategy was still efficient for reaching the solution with the two remaining numerical values.

A solvability judgment task requested participants to tell whether problems were solvable and to write down the solution of the solvable problems. Our aim was to show that semantic constraints associated with cardinal situations would lead adult participants to evoke an encoding incompatible with the 1-step strategy, leading them to fail to find the solution to the problems. We expected them to incorrectly dismiss a perfectly valid solving strategy, erroneously labeling a 1-step subtraction problem as unsolvable. We also expected that succeeding in solving cardinal problems would require an extra representational step, since the initial representation favors the use of the 3-step strategy. We assumed that the construction of a new representation would be costly and time consuming. Therefore, our predictions were twofold: first, we hypothesized that participants would find the solution to ordinal solvable problems more often than to cardinal solvable problems

because of the conflict between a cardinal representation and the 1-step strategy. Second, when correctly solved, cardinal problems would require a significantly longer response time than correctly solved ordinal problems, because of the extra step needed to build a new representation compatible with the 1-step strategy.

Method

Participants. This experiment was conducted online, on the survey platform Qualtrics. Survey link was sent through social networks only. The survey was kept open for one week, after which a total of 89 adults had participated on a voluntary basis (50 women and 39 men, $M = 32.1$ years, $SD = 13.4$ years). All participants spoke French fluently and none had previously participated in any similar experiment. Because part of the analyses were performed on response times, we removed 15 participants who either mentioned taking a break during the test or who answered at least one of the questions in less than 5 seconds (which meant they either mis-clicked or did not take the time to read the problem). The analyses bear on the 74 remaining participants (44 women and 30 men, $M = 33.8$ years, $SD = 13.4$ years).

Materials and procedure. The problems used in this experiment were similar to the ones in previous experiments, except for the value of *Part 1* (see Fig. 1) that was removed from the statements so that the 3-step strategy could not be used anymore. Consequently, the only way to solve the problems was to resort to the 1-step strategy, which required using the remaining values of *Whole 1* and of the *Difference* (see Fig. 1). Table 5 presents 6 examples of such problems (3 cardinal and 3 ordinal problem statements) created from the ones used in the previous experiments by removing the x value corresponding to *Part 1*. Ordinal problems were 333.5 characters long on average ($SD = 38.37$) and cardinal problems were 304 characters long on average ($SD = 44.94$). This length difference was not statistically significant ($t(10) = 1.18$, $p = .26$, paired t-test).

Table 5: Example of target problems used in the study. Changes introduced from the problems in experiments 1 to 3 are italicized in the table for the sake of clarity, but they were not made apparent in the experiment. Translated from French.

Cardinal target problems	Ordinal target problems
Paul has <i>a certain amount of</i> red marbles. He also has blue marbles. In total, Paul has 14 marbles. Jolene has as many blue marbles as Paul, and some green marbles. She has 2 green marbles less than Paul has red marbles. How many marbles does Jolene have?	Sofia travelled <i>for a certain time</i> . Her trip started during the day. Sofia arrived at 14 h. Fred left at the same time as Sofia. Fred's trip lasted 2 hours less than Sofia's. What time was it when Fred arrived?

WHAT WE COUNT DICTATES HOW WE COUNT

In the store, Antoine wants to buy a ruler *costing a certain price*.
He also wants a notebook.
In total, that will cost him 14 dollars.
Julie wants to buy the same notebook as Antoine, and an eraser.
The eraser costs 2 dollars less than the ruler.
How much will Julie have to pay?

Tom takes a Russian dictionary weighing *a certain weight*.
He also takes a Spanish dictionary.
In total, he is carrying 14 kilograms of books.
Lola takes Tom's Spanish dictionary and a German dictionary.
The German dictionary weighs 2 kilograms less than the Russian dictionary.
In total, how many kilograms is Lola carrying?

Slouchy Smurf is *a certain height*.
He climbs on a table.
He now attains the height of 14 centimeters.
Grouchy Smurf climbs on the same table as Slouchy Smurf.
Grouchy Smurf is 2 centimeters shorter than Slouchy Smurf.
What height does Grouchy Smurf attain when he climbs on the table?

Katherine took the elevator and went up *a certain number of floors*.
She left from the floor where the gym is.
She arrived to the 14th floor.
Yohan also took the elevator from the floor where the gym is.
He went up 2 floors less than Katherine.
What floor did Yohan arrive to?

Although our predictions only regarded solvable problems, we also included unsolvable fillers in the materials, so that not every problem had a solution. Among those fillers the value of *Part 1* was preserved, and the value of *Whole 1* was removed, which made the problems unsolvable with either solving strategy. Thus, an equal number of fillers was introduced to achieve a uniform distribution of solvable/unsolvable answers. Problem order and numerical values were randomized between participants. On the first page of the online experiment, the following instructions were written:

You will find an arithmetic problem on each page of this survey. Your task is to identify which problems can be solved and to indicate for each of them the operation you used to solve it, as well as the solution you found. Be careful: some of the problems cannot be solved with the available information, thus your answer in such cases should be 'it is not possible to find the solution'. This is not a speed test: take your time to read and understand each of these problems. *Translated from French.*

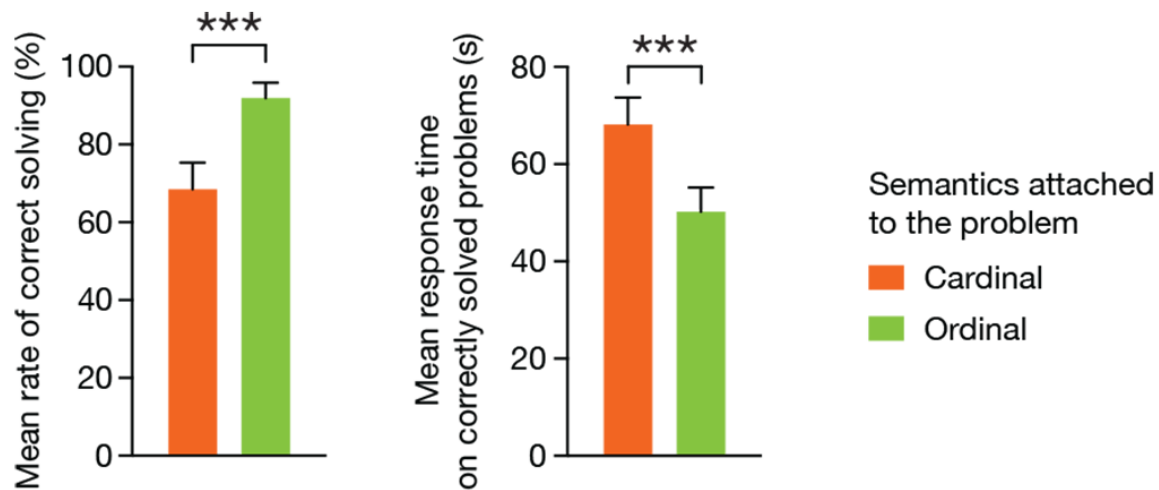
On each page of the survey, a problem was displayed with the following question below it "Given the data provided, is it possible to find the solution?" and two buttons "Yes" and "No". When the participants pressed "Yes", two new questions appeared, asking them to indicate respectively the operation needed to solve the problem and the result of the operation. Participants used the keyboard to write down their answers. After participants answered all 12 problems, a new page was displayed asking them for their gender, date of birth, and whether they made any breaks during the completion of the experiment. Participants all completed the task in less than an hour.

As previously stated, our first prediction was that participants would perform better on solvable problems with ordinal quantities compared to solvable problems with cardinal quantities. Indeed, we believed that problems whose spontaneous representation was associated with a 3-step strategy would often lead participants to ignore

the 1-step strategy, due to cardinal representations being incompatible with the shortest strategy. Our second prediction regarded the cardinal problems that were correctly solved by the participants despite the conflict between a cardinal encoding and the 1-step strategy. We hypothesized that higher response times would be recorded on successfully solved cardinal problems compared to successfully solved ordinal problems, due to participants needing additional time to overcome their initial encoding of the situations and build a new representation, compatible with the 1-step strategy.

Results

The dependent variable was the percentage of correct answers on solvable problems. Ordinal solvable problems were successfully solved in 91.9% of the trials, and cardinal solvable problems in 68.5% of the trials (see left graph of Fig. 7). A paired t-test was performed on participants' mean rate of success for cardinal and ordinal problems and showed that the difference was statistically significant ($t(73) = 6.38, p < .001, d = 0.97$), therefore supporting our first hypothesis.



*Fig. 7: Mean rate of correct resolution (left) and mean response time on correctly solved problems (right) depending on the semantic nature of the quantities used in the problems. Vertical bars denote .95 confidence intervals. *** $p < .001$ (paired t-test).*

Response times on correctly solved cardinal and ordinal problems were then compared in order to test our prediction that accessing the correct 1-step strategy on problems inducing a cardinal encoding incompatible with this strategy would require higher response times than it would on problems evoking an ordinal representation. On average, participants took 68.7 seconds to successfully solve cardinal problems, and 49.8 seconds for ordinal problems (see right graph of Fig. 7). Because we only considered the response times for correctly solved problems, the number of measures per participant could vary from 0 to 6, so we resorted to a mixed model analysis instead of a repeated measures ANOVA. We removed 4 participants who did not manage to correctly solve at least one

cardinal and one ordinal problems, since no comparison could be made between their response times in both conditions. A linear mixed model with participants as a random factor and problem type (cardinal versus ordinal) as a fixed effect showed that the difference between cardinal and ordinal problems had a significant effect on response times of successfully solved problems ($F(1,69) = 20.38, p < .001$), thus supporting our second hypothesis.

Discussion

As predicted, it was more difficult for participants to use the 1-step strategy for cardinal than for ordinal problems, despite the 1-step strategy being the only remaining possibility to find the solution. Besides, as hypothesized, when participants overcame the difficulty and found the solution to a cardinal problem, it required extra processing time, presumably attributable to the construction of a new representation of the situation.

While the fourth experiment showed that the 1-step strategy is more frequently used on ordinal than on cardinal problems, this fifth experiment showed that this effect is not the result of a mere preference but, instead, seems to be the consequence of strong limitations imposed by the type of quantity used. In fact, this effect was so pervasive that in many cases adult participants failed to see that these one-step subtraction problems could be solved at all. Yet, not all the participants failed, and some of them even managed to use the 1-step strategy in certain cases. However, in order to overcome the constraints imposed by their world knowledge about the problem's quantities and use a conflicting solving strategy, the participants had to discard their initial representation and construct a new encoding closer to the problem's mathematical structure. The existence of such a recoding step, akin to a re-representation process (Vicente, Orrantia, & Verschaffel, 2007), was supported by the longer response times required on correctly solved cardinal problems.

Once participants realize that the 3-step strategy cannot be used given the available information, they might be tempted to discard these problems as unsolvable and move on. In the 6th experiment, we went a step further and provided participants with a potential solution to the problems, one that they would not usually consider. By giving them such a clue, we were able to assess their difficulty to construct an alternate encoding of the situation even when directly incited to do so.

Experiment 6

In this experiment, we provided participants with the 1-step solving strategy of each problem and asked them to directly evaluate its validity. Because of the high failure rates on cardinal problems in Experiment 5, we tested the stronger hypothesis that providing the solving strategy would not be enough to systematically foster an appropriate encoding of the situation by the participants. We hypothesized that even when explicitly presented

with the solution, participants would reject it more often in the cardinal condition than in the ordinal one since their encoding of cardinal problems would conflict with their solution (i.e. promoting an unusable 3-step strategy instead of the 1-step strategy). Because of the need to overcome the initial representation, we hypothesized that the correct identification of the solution would require more time for cardinal than for ordinal problems. This experimental paradigm resembled the one used in a recent study we conducted among expert mathematicians (Gros, Sander, & Thibaut, 2019). However, in the present experiment, participants could take as long as they wished to complete the task, whereas in Gros et al. (2019), participants were explicitly told to solve the problems as fast as possible, with the intended purpose of increasing their error rates. The absence of time constraint was meant to give participants the opportunity to read the problems until they were certain of their decision and to engage in a recoding of their initial representation if need be.

Method

Participants. This experiment was conducted online, on the survey platform Qualtrics. Survey link was sent through a Parisian mailing list for cognitive science experiments. We kept the survey open for one week, after which a total of 223 adults had participated on a voluntary basis. The increase in sample size is attributed to a response rate above our expectations on the mailing list for voluntary participants. All spoke French fluently and none had previously participated in any similar experiment. Among them, 27 were removed from the analysis because they either took a break during the test or answered at least one of the questions in less than 5 seconds (which suggested they did not take the time to read the problem). The analyses were performed on the remaining 196 participants (135 women and 88 men, $M = 34.5$ years, $SD = 14.8$ years).

Materials and procedure. The only difference between the present experiment and the previous one was the fact that a solution was proposed. Instead of having the participants solve the problems themselves, a solution was proposed for each problem, and they were asked to judge whether the provided solution was valid or whether the problem was unsolvable. For every problem, the question “Given the data provided, is it possible to find the solution?” was displayed. Two choices appeared below: (a) “No, we do not have enough information to solve this problem.” and (b) “Yes: *numerical value 1 – numerical value 2 = result. Sentence presenting the result*”. For instance, on one of the elevator problems, the option (b) was: “Yes: $11 - 2 = 9$. Karin arrives at the 9th floor.” Participants all completed the task in less than an hour.

Results

As in Experiment 5, we first analyzed the ratio of correct answers on solvable problems depending on the type of

quantities used. The left graph of Fig. 8 shows that, among the target problems, the cardinal ones had a lower success rate (63.6%) than the ordinal ones (88.4%). A paired t-test performed on the participants' mean rate of success confirmed that this difference was significant ($t(195) = 9.25, p < .001, d = 0.87$).

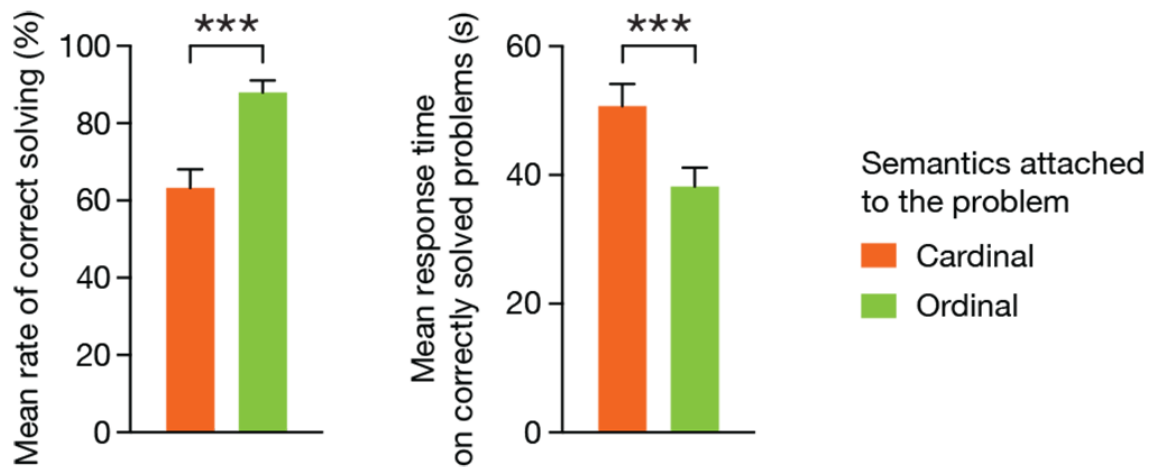


Fig. 8: Mean rate of correct resolution (left) and mean response time on correctly solved problems (right) depending on the semantic nature of the quantities used in the problems. Vertical bars denote .95 confidence intervals. *** $p < .001$.

In order to assess the validity of our second hypothesis, we analyzed the response times of correct answers on the target solvable problems. The 26 participants who did not manage to correctly respond to at least one cardinal and one ordinal problems were removed from this analysis, since no comparison could be made between their response times in both conditions. The right graph of Fig. 8 shows that providing a correct answer required a shorter response time for ordinal (38.6 seconds) than for cardinal problems (51.4 seconds). A linear mixed-model with participants as a random effect and the cardinal versus ordinal distinction as a fixed factor confirmed that the effect was statistically significant ($F(1,169) = 30.28, p < .001$), supporting the second hypothesis.

Discussion

This experiment, involving a solution validity assessment task, supported the effects observed in the previous one, involving a solution discovery task. The analyses indicated that even when the correct solution was provided, it was more difficult for them to accept it when it was not compatible with the initial encoding of the problem. Despite the problems being solvable with a mere subtraction, participants went so far as to reject the correct solution and dismiss the problems as “unsolvable”. Furthermore, overcoming this difficulty required more time, thus supporting our prediction of the need for an extra processing step when faced with an inapt representation. These results suggest that the encoding effects identified in the 5th experiment are not restricted to the elaboration of a solving strategy, but also to the evaluation of its validity. This experiment provides additional evidence that

the cardinal versus ordinal distinction constrains the encoding of problems, since even when no solving strategy had to be produced by the participants, their interpretation precluded them from considering the given solution – albeit a single subtraction – as an acceptable one.

General discussion

The pervasive influence of the cardinal-ordinal distinction

Taken together, the present six experiments shed light upon the foundational part played by the cardinal versus ordinal distinction in adults' reasoning about numerical situations. The first experiment confirmed the relevance of the selected quantities, as the distinction between cardinal and ordinal quantities drove adults' sorting patterns. Experiments 2 and 3 demonstrated that the encoding difference between cardinal and ordinal problems impacts participants' success in perceiving problems as analogous, even when explicitly instructed to directly compare two problems. The fourth experiment's findings were twofold. First, it showed that the distinction between cardinal and ordinal problems influences adults' choice of a solving strategy. Second, it demonstrated that it is possible to manipulate a problem's semantics by presenting cardinal quantities in a context emphasizing the ordinality of the problem's values. The changes that were introduced to make the cardinal problems more ordinal had a significant effect on participants' strategy choice, thus bolstering the importance of the cardinal versus ordinal distinction and strengthening the view that the encoding difference observed depended on the semantic dimensions that were manipulated in the problems. The fifth experiment showed that even when only one solution was available, adult participants have difficulties to find it when it is not compatible with what is assumed to be their spontaneous encoding of the problems. They were more likely to erroneously judge that cardinal problems cannot be solved and, when they did find the solution to a cardinal problem, it nonetheless required a longer reasoning time. Finally, the sixth experiment demonstrated that presenting the participants with a potential solution does not suppress the effects observed in Experiment 5, as participants continued to struggle to identify the solution of cardinal problems.

The fact that the distinction between cardinal and ordinal quantifications could exert a pervasive effect on adults' grasp of numerical situations illustrates the foundational nature of this distinction for the human mind. Despite a growing body of research on the comparative development of cardinality and ordinality, little is known regarding adults' ability to alternatively perceive the cardinality or the ordinality of the numerical situations they encounter, depending on the daily-life knowledge imbued in these situations. Our understanding of mathematics is deeply rooted into our understanding of the world (Fischbein, 1987; Hofstadter & Sander, 2013; Lakoff &

Núñez, 2000) and, as a result, we tend to apply real-life constraints to abstract mathematical concepts. Here, we showed that our perception of ordinality and cardinality in scenes of our daily lives has a profound effect on our ability to engage in mathematical reasoning in these situations.

Implications for mathematical word problem solving

Although converging works have shown that slight modifications in the wording of otherwise structurally identical mathematical word problems could result in significant performance disparities (Carpenter & Moser, 1982; Cummins, Kintsch, Reusser, & Weimer, 1988; De Corte, Verschaffel, & De Win, 1985; Hudson, 1983; Nesher, Greeno, & Riley, 1982; Scheibling-Sève & Sander 2020; Squire & Bryant, 2002; Thevenot & Oakhill, 2005), the identification of the underlying reasoning processes accounting for such variations has remained a debated issue to this day. One possible explanation can be found in the schema theory (Kintsch, 1988; Kintsch & Greeno, 1985; Rumelhart, 1980; Schank, 1975; Schank & Abelson, 1977). However, a number of shortcomings have been shown to hinder the schema theory's explanatory power of the interpretative effects that seem to be central in mathematical word problem solving (Devidal, Fayol, & Barrouillet, 1997; Fayol, Abdi, & Gombert, 1987; Thevenot, 2010, 2017; Thevenot, Devidal, Barrouillet, & Fayol, 2007). In our case, Experiment 4's results showed that participants use different solving strategies to solve isomorphic problems, depending on the type of quantities these problems mention. In the seminal description of the schema theory, Kintsch and Greeno (1985) indicate that the activation of a schema to solve a problem is done based on the propositional structure of the problem statement. In other words, the relevant entities are the numerical values and the relations explicitly described in the problems. The model proposed by Kintsch and Greeno does not explicitly raise the possibility that the world knowledge evoked by the elements mentioned in a problem may influence solvers' representation of the problems, nor their choice of a solving strategy. Thus, those results could not have been predicted by the schema theory.

Since 1990, an alternative approach has emerged from Reusser's critical observations on the schema theory: the Situation-Problem-Solver, a model introducing the idea that the text of the problem statement is first translated into a situation model of the situation, before being "mathematized" into a problem model and finally translated into a solving algorithm (Reusser, 1990; Staub & Reusser, 1995). Yet, it has been argued that by relying on the notion of situation model, the Situation-Problem-Solver approach neglects some interpretative effects at play in mathematical word problem solving. Indeed, this model does not directly include the idea that different models of a given problem statement can be built, depending on the solver's viewpoint (Gros, Thibaut, & Sander, 2020). More specifically, this approach struggles to account for the idea that the initial problem representation

constructed by the solvers may vary depending on the state of their general, non-mathematical knowledge about the elements present in the problems. Moreover, the notion that participants may overcome their initial encoding of a problem and construct a new representation also falls outside of the Situation-Problem-Solver's scope. However, in Experiments 5 and 6, participants were shown to engage in a semantic recoding process, akin to re-representation (Vicente et al., 2007), when their initial representation of the problem led to a dead end.

Thus, the schema theory and the Situation-Problem-Solver model both struggle to account for the results of our six experiments. On the other hand, those results are well within the scope of the new Semantic Congruence (SECO) model, which provides a conceptual framework for arithmetic word problem solving (Gros et al., 2020). The SECO model is meant to account for the influence of world semantics (the non-mathematical world knowledge one has about the entities described in a problem statement) on the initial encoding of a word problem into a specific interpreted structure (see Fig. 9). It proposes that when reading a problem statement, most solvers resort to their world semantics as well as to their mathematical semantics (the solver's mathematical knowledge applicable to the problem statement) to construct a representation of the described situation (an interpreted structure), which can then be translated into specific solving strategies. When no relevant solving strategy can be specified from the interpreted structure, then solvers may attempt to engage in semantic recoding by using mathematical knowledge to construct a new representation of the problem at hand. When performed successfully, this semantic recoding makes it possible to use different solving strategies compatible with the new constructed representation (Gros et al., 2020).

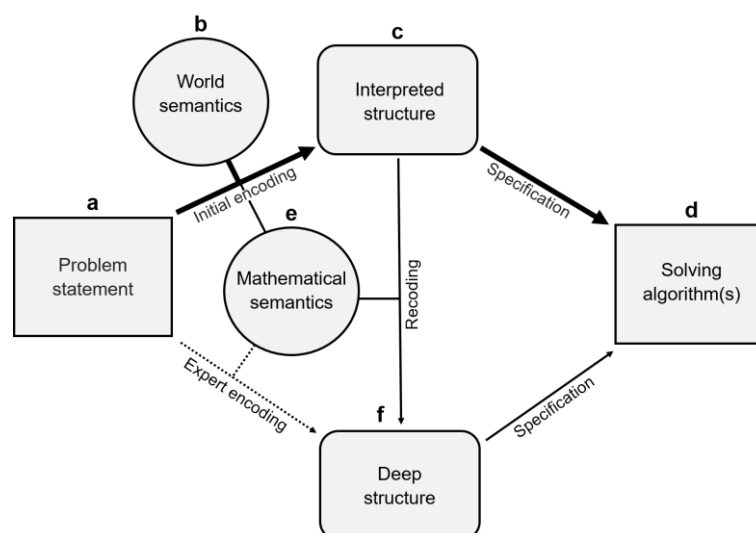


Fig. 9: Structure of the SECO model. Reprinted from “Semantic congruence in arithmetic: A new model for word problem solving,” by H. Gros, J.-P. Thibaut, & E. Sander, 2020, *Educational Psychologist*, 55, pp. 69-87.

For instance, in the case of the problems used in this study, SECO would predict that a cardinal problem featuring marbles will activate the solver's world knowledge about marbles (e.g. the fact that marbles are

unordered entities that are usually grouped in no specific order) which will lead to the encoding of a cardinal interpreted structure, which then leads to the use of a 3-step strategy (see Experiment 4). However, when a numerical value is removed from the problem and the 3-step strategy cannot be used (see Experiments 5 and 6), participants need to discard the irrelevant surface features of the problem and engage in a semantic recoding process focused on the problem's mathematical properties (the notion that the difference between two sets sharing a common part is equal to the difference between their disjoint parts). If performed successfully, this recoding makes it possible for the participants to understand the relevance of the 1-step strategy to solve cardinal problems. SECO suggests that engaging in such a recoding is cognitively costly, which explains the increased errors and reaction times on cardinal problems in these experiments. The results reported in this study thus support the relevance of the SECO model to account for arithmetic word problem solving, compared to previous approaches in the field.

Educational perspectives

The range of content effects we have put forward in this paper showcase the fact that even for problems involving only elementary arithmetic operations, supposedly acquired in primary school, humans hardly manage to completely disregard context, and their performances in problem solving suffer from this shortcoming. However, the influence of world knowledge on mathematical word problem solving is not surprising considering that, as suggested by the *kind world hypothesis*, it leads to the making of inferences which are correct most of the time, since the surface features of situations are highly correlated with deeper principles (Bassok et al., 2008; Blessing & Ross, 1996; Gentner & Medina, 1998; Goldstone, 1994; Trench & Minervino, 2015). Consequently, people tend to rely on those superficial cues which help them find the solution. This is especially true for novices, who predominantly use surface features to recall previous experiences, while experts may be more apt to perceive the structure of the situations they encounter (Raynal, Clément, & Sander, 2020). Difficulties arise when the world knowledge evoked by a problem and its deep structure are semantically incongruent. Unexperienced learners are then bound to struggle, and sometimes even fail. This is especially problematic since mathematics education does not usually control for content effects (Bassok et al., 1998; Gvozdic & Sander, 2018; Lee, DeWolf, Bassok, & Holyoak, 2016), which is partly due to mathematics being primarily considered the realm of abstraction (Davis, Hersh, & Marchisotto, 2011; Russell, 1903).

Although arithmetic word problems are a central part of mathematics education and teachers are usually encouraged to provide real-world examples to illustrate the concepts being taught (e.g. Richland, Stigler, & Holyoak, 2012; Rivet & Krajcik, 2008), the use of concrete examples to teach new concepts has also been shown

to have a detrimental effect on transfer (Day, Motz, & Goldstone, 2015; Goldstone & Sakamoto, 2003; Son & Goldstone, 2009). In line with current efforts to develop new teaching methods aimed at overcoming the deleterious influence of content effects (e.g. Gvozdic & Sander, 2019), our findings on the perception of cardinality and ordinality in concrete situations may hold value to design new teaching sequences. Just as *concreteness fading* is proposed as a tool to improve transfer by resorting to increasingly abstract examples (Fyfe, McNeil, Son, & Goldstone, 2014; Fyfe & Nathan, 2019), it may be a promising route to develop a *semantic congruence fading* process using increasingly incongruent examples (Gros et al., 2020). In the case of the problems used in the current study, starting with teaching the 1-step strategy on ordinal problems, then moving to hybrid problems and finally to cardinal problems may help learners acquire a better understanding of this strategy, and they may consequently learn to use it in any situation, regardless of the semantics conveyed by the problem statement. We believe that building on the semantic congruence framework to develop such training programs may help learners develop a richer understanding of cardinality and ordinality as well as a more flexible approach to the encoding of numerical situations in general.

References

- Bagnoud, J., Burra, N., Castel, C., Oakhill, J., & Thevenot, C. (2018). Arithmetic word problems describing discrete quantities: EEG evidence for the construction of a situation model. *Acta psychologica*, 190, 116-121.
- Bassok, M. (2001). Semantic alignments in mathematical word problems. *The analogical mind: Perspectives from cognitive science*, 401-433.
- Bassok, M., Chase, V. M., & Martin, S. A. (1998). Adding apples and oranges: Alignment of semantic and formal knowledge. *Cognitive Psychology*, 35, 99-134.
- Bassok, M., Pedigo, S. F., & Oskarsson, A. (2008). Priming addition facts with semantic relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(2), 343-352.
- Bassok, M., Wu, L. L., & Olseth, K. L. (1995). Judging a book by its cover: Interpretative effects of content on problem-solving transfer. *Memory and Cognition*, 23, 354-367.
- Bermejo, V. (1996). Cardinality development and counting. *Developmental Psychology*, 32(2), 263-268.
- Blessing, S. B., & Ross, B. H. (1996). Content effects in problem categorization and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(3), 792.
- Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: evidence for a mental time line. *Neuroscience & Biobehavioral Reviews*, 36(10), 2257-2273.
- Boroditsky, L. (2011). How language shapes thought. *Scientific American*, 304(2), 62-65.
- Carpenter, T. P., Lindquist, M. M., Brown, C. A., Kouba, V. L., Silver, E. A., & Swafford, J. O. (1988). Results of the fourth NAEP Assessment of Mathematics: Trends and conclusions. *Arithmetic Teacher*, 36(4), 38-41.
- Carpenter, T. P., & Moser, J. M. (1982). The development of addition and subtraction problem-solving skills. In T. P. Carpenter, J. M. Moser & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective*, (pp. 9-24). Hillsdale, NJ: Erlbaum.
- Casasanto, D. (2008). Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Language learning*, 58, 63-79.
- Chen, Z., Xu, M., Shang, D., Peng, G., & Luo, B. (2014). Distinct representations of symbolic ordinality and quantity: evidence from neuropsychological investigations in a Chinese patient with Gerstmann's syndrome. *Brain and cognition*, 88, 14-20.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems

- by experts and novices. *Cognitive Science*, 5, 121-152.
- Clément, E., & Richard, J.-F. (1997). Knowledge of domain effects in problem representation: The case of Tower of Hanoi isomorphs. *Thinking & Reasoning*, 3(2), 133-157.
- Colomé, À., & Noël, M. P. (2012). One first? Acquisition of the cardinal and ordinal uses of numbers in preschoolers. *Journal of Experimental Child Psychology*, 113(2), 233-247.
- Condry, K. F., & Spelke, E. S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology: General*, 137(1), 22-38.
- Cook, J. L., & Rieser, J. J. (2005). Finding the critical facts: Children's visual scan patterns when solving story problems that contain irrelevant information. *Journal of educational psychology*, 97(2), 224.
- Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive psychology*, 20(4), 405-438.
- Dantzig, T. (1945). *Number the language of science. A critical survey written for the cultured non-mathematician*. New York, NY: The Macmillan Company.
- Daroczy, G., Wolska, M., Meurers, W. D., & Nuerk, H. C. (2015). Word problems: a review of linguistic and numerical factors contributing to their difficulty. *Frontiers in psychology*, 6, 348.
- Davis, P., Hersh, R., & Marchisotto, E. A. (2011). *The mathematical experience*. Springer Science & Business Media.
- Day, S. B., Motz, B. A., & Goldstone, R. L. (2015). The cognitive costs of context: The effects of concreteness and immersiveness in instructional examples. *Frontiers in Psychology*, 6, 1876.
- De Corte, E., & Verschaffel, L. (1985). Beginning first graders' initial representation of arithmetic word problems. *The Journal of Mathematical Behavior*, 4, 3-21.
- De Corte E., Verschaffel L. (1987). The effect of semantic structure on first graders' strategies for solving addition and subtraction word problems. *Journal of Research in Mathematics Education*, 18, 363–381.
- De Corte, E., Verschaffel, L., & De Win, L. (1985). Influence of rewording verbal problems on children's problem representations and solution. *Journal of Educational Psychology*, 77, 460-470.
- Delazer, M., & Butterworth, B. (1997). A dissociation of number meanings. *Cognitive Neuropsychology*, 14(4), 613-636.
- Devidal, M., Fayol, M., & Barrouillet, P. (1997). Stratégies de lecture et résolution de problèmes arithmétiques. *L'Année Psychologique*, 97(1), 9-31.

- Droit-Volet, S., & Coull, J. (2015). The developmental emergence of the mental time-line: spatial and numerical distortion of time judgement. *PloS one*, 10(7), e0130465.
- Fayol, M., Abdi, H., & Gombert, J. E. (1987). Arithmetic problems formulation and working memory load. *Cognition and Instruction*, 4(3), 187-202.
- Fischbein, H. (1987). *Intuition in science and mathematics: An educational approach*. (Vol. 5). Springer Science & Business Media.
- Fischer, F. E., & Beckey, R. D. (1990). Beginning kindergarteners' perception of number. *Perceptual and Motor Skills*, 70(2), 419-425.
- Frege, G. (1980). *The Foundations of Arithmetic* [1884], transl. J. Austin. Evanston, IL: Northwestern University Press.
- Fuhrman, O., & Boroditsky, L. (2010). Cross-cultural differences in mental representations of time: Evidence from an implicit nonlinguistic task. *Cognitive Science*, 34(8), 1430-1451.
- Fuson, K. C. (1988). *Children's counting and concepts of number*. New York: Springer-Verlag.
- Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, 26(1), 9-25.
- Fyfe, E. R., & Nathan, M. J. (2019). Making “concreteness fading” more concrete as a theory of instruction for promoting transfer. *Educational Review*, 71(4), 403-422.
- Gamo, S., Sander, E., & Richard, J.-F. (2010). Transfer of strategy use by semantic recoding in arithmetic problem solving. *Learning and Instruction*, 20, 400-410.
- Gelman, R., & Gallistel, C. R. (1978). *The child's concept of number*. Cambridge, MA: Harvard University Press.
- Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition*, 65(2-3), 263-297.
- Goffin, C., & Ansari, D. (2016). Beyond magnitude: Judging ordinality of symbolic number is unrelated to magnitude comparison and independently relates to individual differences in arithmetic. *Cognition*, 150, 68-76.
- Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition*, 52(2), 125-157.
- Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive psychology*, 46(4), 414-466.
- Gros, H., Sander, E., & Thibaut, J. -P. (2019). When masters of abstraction run into a concrete wall: Experts

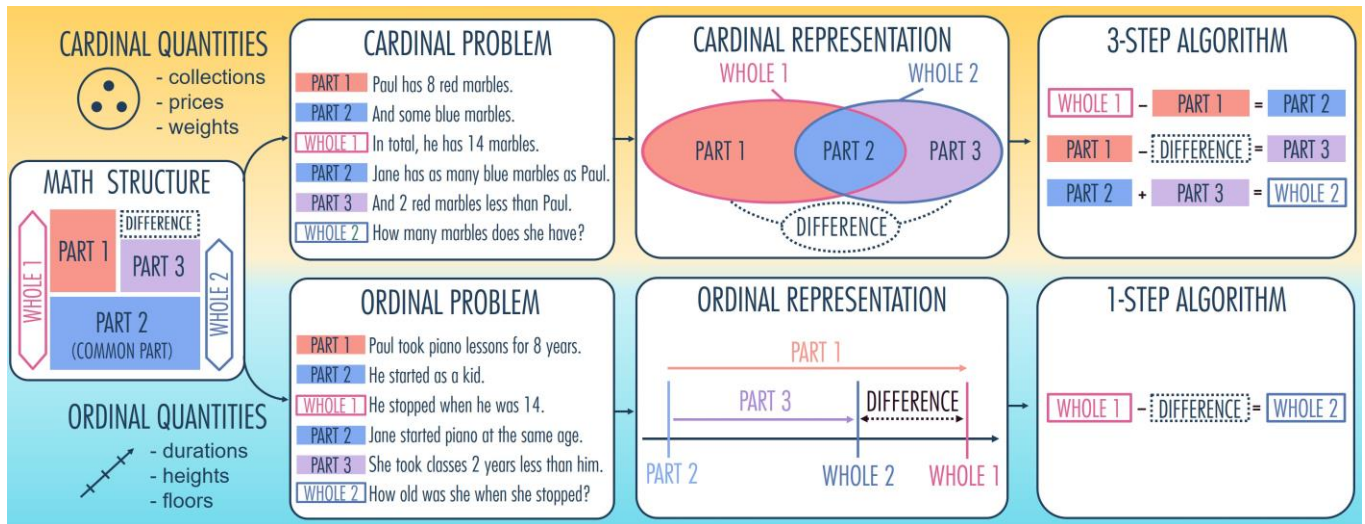
- failing arithmetic word problems. *Psychonomic bulletin & review*, 26(5), 1738-1746.
- [dataset] Gros, H., Thibaut, J.-P., & Sander, E. *What we count dictates how we count: A tale of two encodings*, OSF, 2020. https://osf.io/kz6gh/?view_only=2f3fb6b910844e238ae58fef3c61168a
- Gros, H., Thibaut, J. -P., & Sander, E. (2017). The nature of quantities influences the representation of arithmetic problems: Evidence from drawings and solving procedures in children and adults. In R. Granger, U. Hahn, & R. Sutton (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp 439–444). Austin, TX: Cognitive Science Society.
- Gros, H., Thibaut, J. -P., & Sander, E. (2020) Semantic Congruence in Arithmetic: A New Model for Word Problem Solving, *Educational Psychologist*, 55(2), 69-87.
- Gvozdic, K., & Sander, E. (2018). When intuitive conceptions overshadow pedagogical content knowledge: Teachers’ conceptions of students’ arithmetic word problem solving strategies. *Educational Studies in Mathematics*, 98, 157–175.
- Gvozdic, K., & Sander, E. (2020). Learning to be an opportunistic word problem solver: Going beyond informal solving strategies. *ZDM Mathematics Education*, 52, 111-123.
- Hembree, R. (1992). Experiments and relational studies in problem solving: A meta-analysis. *Journal for Research in Mathematics Education*, 242-273.
- Hofstadter, D. R., & Sander, E. (2013). *Surfaces and essences: Analogy as the fuel and fire of thinking*. New York: Basic Books.
- Hood, B. M. (1998). Gravity does rule for falling events. *Developmental Science*, 1(1), 59-63.
- Hood, B. M., Santos, L., & Fieselman, S. (2000). Two-year-olds’ naïve predictions for horizontal trajectories. *Developmental Science*, 3(3), 328-332.
- Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development*, Vol. 54, No. 1, 84-90.
- Izard, V., Streri, A., & Spelke, E. S. (2014). Toward exact number: Young children use one-to-one correspondence to measure set identity but not numerical equality. *Cognitive Psychology*, 72, 27-53.
- Johnson, P. E., Moen, J. B., & Thompson, W. B. (1988) Garden Path Errors in Diagnostic Reasoning. In: Bolc L., Coombs M.J. (eds) *Expert System Applications. Symbolic Computation (Artificial Intelligence)*. Springer, Berlin, Heidelberg.
- Kim, I. K., & Spelke, E. S. (1992). Infants' sensitivity to effects of gravity on visible object motion. *Journal of Experimental Psychology: Human Perception and Performance*, 18(2), 385.

- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological review*, 95(2), 163.
- Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological review*, 92(1), 109.
- Kouba, V. L., & Wearne, D. (2000). Whole number properties and operations. In E. A. Silver & P. A. Kenney (Eds.), *Results from the Seventh Mathematics Assessment of the National Assessment of Educational Progress* (pp. 141–161). Reston, VA: National Council of Teachers of Mathematics.
- Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: how the embodied mind brings mathematics into being* (Nachdr.). New York, NY: Basic Books.
- Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, 105(2), 395-438.
- Lee, H. S., DeWolf, M., Bassok, M., & Holyoak, K. J. (2016). Conceptual and procedural distinctions between fractions and decimals: A cross-national comparison. *Cognition*, 147, 57-69.
- Low, R., & Over, R. (1989). Detection of missing and irrelevant information within algebraic story problems. *British Journal of Educational Psychology*, 59, 296–305.
- Lyons, I. M., & Beilock, S. L. (2013). Ordinality and the nature of symbolic numbers. *Journal of Neuroscience*, 33(43), 17052-17061.
- Lyons, I. M., Vogel, S. E., & Ansari, D. (2016). On the ordinality of numbers: A review of neural and behavioral studies. In *Progress in brain research*, 227, 187-221.
- Matejko, A. A., Hutchison, J. E., & Ansari, D. (2019). Developmental specialization of the left intraparietal sulcus for symbolic ordinal processing. *Cortex*, 114, 41-53.
- Meyer, C., Barbiers, S., & Weerman, F. (2016). Order and Ordinality: The Acquisition of Cardinals and Ordinals in Dutch. In *Boston University Conference on Language Development (BUCLD) 40*. MA: Cascadilla Press.
- Miller, K., Major, S. M., Shu, H., & Zhang, H. (2000). Ordinal knowledge: Number names and number concepts in Chinese and English. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 54(2), 129.
- Miller, S. E., Marcovitch, S., Boseovski, J. J., & Lewkowicz, D. J. (2015). Young children's ability to use ordinal labels in a spatial search task. *Merrill-Palmer Quarterly*, 61(3), 345-361.
- Nesher, P., Greeno, J. G., & Riley, M. S. (1982). The development of semantic categories for addition and

- subtraction. *Educational Studies in Mathematics*, 13(4), 373–394.
- Ng, J., Lee, K., & Khng, K. H. (2017). Irrelevant information in math problems need not be inhibited: Students might just need to spot them. *Learning and Individual Differences*, 60, 46-55.
- Northcote, M. T., & Marshall, L. (2016). What mathematics calculations do adults do in their everyday lives?: Part 1 of a report on the Everyday Mathematics Project. *Australian Primary Mathematics Classroom*, 21 (2), 8-17.
- Piaget, J. (1965). *The child's conception of number*. New York: W. W. Norton. (Original work published 1941).
- Raynal, L., Clément, E., & Sander, E. (2020). Are Superficially Dissimilar Analogs better retrieved than Superficially Similar Disanalogs?. *Acta Psychologica*, 203, 102989.
- Reusser, K. (1990). From text to situation to equation: Cognitive simulation of understanding and solving mathematical word problems. In H. Mandl, E. De Corte, N. Bennet & H.F. Friedrich (Eds), *Learning and Instruction, European Research in an International Context*, Vol. II. New York: Pergamon Press.
- Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). *Teaching the conceptual structure of mathematics*. *Educational Psychologist*, 47(3), 189-203.
- Rivet, A. E., & Krajcik, J. S. (2008). Contextualizing instruction: Leveraging students' prior knowledge and experiences to foster understanding of middle school science. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 45(1), 79-100.
- Rubinsten, O., Dana, S., Lavro, D., & Berger, A. (2013). Processing ordinality and quantity: ERP evidence of separate mechanisms. *Brain and Cognition*, 82(2), 201-212.
- Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33-58). Hillsdale, NJ: Erlbaum.
- Russell, B. (1903). *Principles of mathematics*. Cambridge, UK: Cambridge University Press.
- Russell, B. (1919). *Introduction to mathematical philosophy*. London: Allen & Unwin.
- Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, 108(3), 662-674.
- Sarnecka, B. W., & Lee, M. D. (2009). Levels of number knowledge during early childhood. *Journal of Experimental Child Psychology*, 103(3), 325-337.
- Sasanguie, D., Lyons, I. M., De Smedt, B., & Reynvoet, B. (2017). Unpacking symbolic number comparison

- and its relation with arithmetic in adults. *Cognition*, 165, 26-38.
- Schank, R. C. (1975). The role of memory in language processing. In: C. N. Cofer (ed), *The Structure of Human Memory*. San Francisco. (1975): p. 162-189
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans and understanding: An inquiry into human knowledge structures*. New York: Hillsdale.
- Scheibling-Sève, C., Pasquinelli, E., & Sander, E. (2020). Assessing conceptual knowledge through solving arithmetic word problems. *Educational Studies in Mathematics*, online first.
- Schoenfeld, A. (2009). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. *Colección Digital Eudoxus*, (7).
- Simon, T. J. (1997). Reconceptualizing the origins of number knowledge: A “non-numerical” account. *Cognitive Development*, 12(3), 349-372.
- Son, J. Y., & Goldstone, R. L. (2009). Contextualization in perspective. *Cognition and Instruction*, 27(1), 51-89.
- Squire, S., & Bryant, P. (2002). From sharing to dividing: Young children’s understanding of division. *Developmental Science*, 5(4), 452–466.
- Stacey, K. (2005). The place of problem solving in contemporary mathematics curriculum documents. *The Journal of Mathematical Behavior*, 24(3-4), 341-350.
- Staub, F. C., & Reusser, K. (1995). The role of presentational structures in understanding and solving mathematical word problems. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension* (pp. 285–305). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Suppes, P. (1972). *Axiomatic set theory*. New York, NY: Dover Publications.
- Suzuki, R., & Shimodaira, H. (2006). Pvcust: An R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, 22(12), 1540-1542.
- Thevenot, C. (2010). Arithmetic word problem solving: Evidence for the construction of a mental model. *Acta psychologica*, 133(1), 90-95.
- Thevenot, C. (2017). Arithmetic word problem solving: the role of prior knowledge. In D. C. GearyD. B. Berch, R. Ochsendorf, K. M. Koepke (Eds.), *Mathematical cognition and learning: Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 47-66). London: Academic Press.
- Thevenot, C., Devidal, M., Barrouillet, P., & Fayol, M. (2007). Why does placing the question before an

- arithmetic word problem improve performance? A situation model account. *The Quarterly Journal of Experimental Psychology*, 60(1), 43-56.
- Thevenot, C., & Oakhill, J. (2005). The strategic use of alternative representations in arithmetic word problem solving. *Quarterly Journal of Experimental Psychology-A*, 58(7), 1311-1323.
- Tomasello, M., & Herron, C. (1988). Down the garden path: Inducing and correcting overgeneralization errors in the foreign language classroom. *Applied Psycholinguistics*, 9(3), 237-246.
- Trench, M., & Minervino, R. A. (2015). The role of surface similarity in analogical retrieval: Bridging the gap between the naturalistic and the experimental traditions. *Cognitive science*, 39(6), 1292-1319.
- Turconi, E., Jemel, B., Rossion, B., & Seron, X. (2004). Electrophysiological evidence for differential processing of numerical quantity and order in humans. *Cognitive Brain Research*, 21(1), 22-38.
- Turconi, E., & Seron, X. (2002). Dissociation between order and quantity meaning in a patient with Gerstmann syndrome. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 38(5), 911-914.
- Verschaffel, L., De Corte, E., & Vierstraete, H. (1999). Upper elementary school pupils' difficulties in modeling and solving nonstandard additive word problems involving ordinal numbers. *Journal for Research in Mathematics Education*, 30(3), 265-285
- Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: a survey. *ZDM*, 52(1), 1-16.
- Vicente, S., Orrantia, J., & Verschaffel, L. (2007). Influence of situational and conceptual rewording on word problem solving. *British Journal of Educational Psychology*, 77(4), 829-848.
- Vondrová, N. (2020). The effect of an irrelevant number and language consistency in a word problem on pupils' achievement and reasoning. *International Journal of Mathematical Education in Science and Technology*, 1-20.
- Wasner, M., Moeller, K., Fischer, M. H., & Nuerk, H. C. (2015). Related but not the same: Ordinality, cardinality and 1-to-1 correspondence in finger-based numerical representations. *Journal of Cognitive Psychology*, 27(4), 426-441.
- Weger, U. W., & Pratt, J. (2008). Time flies like an arrow: Space-time compatibility effects suggest the use of a mental timeline. *Psychonomic Bulletin & Review*, 15(2), 426-430.
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358(6389), 749-750.

Graphical abstract:**Highlights**

- Humans alternatively perceive the cardinality or ordinality of numerical situations
- Arithmetic word problems are encoded based on daily-life non-mathematical knowledge
- Use of collections, prices, or weights fosters set-based, cardinal representations
- Use of durations, heights, or floors fosters axis-based, ordinal representations
- Solving an incongruent problem requires to recode the constructed representation

Keywords:

Arithmetic problem solving; Cardinality; Mathematical reasoning; Mental model; Numerical cognition; Ordinality.