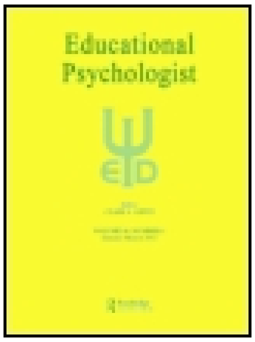

Educational Psychologist

ISSN: 0046-1520 (Print) 1532-6985 (Online) Journal homepage: https://www.tandfonline.com/loi/hedp20

Semantic congruence in arithmetic: A new conceptual model for word problem solving

Hippolyte Gros, Jean-Pierre Thibaut & Emmanuel Sander

EDUCATIONAL PSYCHOLOGIST
https://doi.org/10.1080/00461520.2019.1691004

Semantic congruence in arithmetic: A new conceptual model for word problem solving

Hippolyte Gros[a], Jean-Pierre Thibaut[b], and Emmanuel Sander[c]

[a]Center for Research and Interdisciplinarity, Paris Descartes University, France; [b]LEAD, CNRS UMR 5022, University of Bourgogne Franche-Comté, Dijon, France; [c]Faculty of Psychology and Educational Sciences, University of Geneva, Switzerland

ABSTRACT
Arithmetic problem solving is a crucial part of mathematics education. However, existing problem solving theories do not fully account for the semantic constraints partaking in the encoding and recoding of arithmetic word problems. In this respect, the limitations of the main existing models in the literature are discussed. We then introduce the Semantic Congruence (SECO) model, a theoretical model depicting how world and mathematical semantics interact in the encoding, recoding, and solving of arithmetic word problems. The SECO model's ability to account for emblematic results in educational psychology is scrutinized through six case studies encompassing a wide range of effects observed in previous works. The influence of world semantics on learners' problem representations and solving strategies is put forward, as well as the difficulties arising from semantic incongruence between representations and algorithms. Special attention is given to the recoding of semantically incongruent representations, a crucial step that learners struggle with.

What does it take to solve an arithmetic word problem? It goes without saying that finding the solution requires to be able to read and understand the problem statement, as well as to handle its numerical values and compute the solving algorithm. But is it enough to simply know how to read and count?

Several studies have highlighted robust effects suggesting that solving arithmetic word problems involves processes other than mere procedural ones, that have yet to be accounted for within a unified theory. For instance, Hudson (1983) showed that finding a solution to the problem "There are 5 birds and 3 worms. How many more birds than worms are there?" was considerably more difficult for kindergarteners than answering the question "How many birds won't get a worm?," despite striking similarities between these two situations. Bassok, Wu, and Olseth (1995) showed that after being taught the algorithmic solution of a problem describing objects assigned to people (e.g., computers given to secretaries), participants could more easily transfer it to problems involving objects assigned to people (e.g., prizes given to students), rather than to problems involving different semantic relations, such as problems involving symmetrical sets of people (e.g., doctors "assigned" to other doctors). In a study with primary school pupils, Coquin-Viennot and Moreau (2003) found that to calculate the number of flowers a florist needs in order to give five roses and seven tulips to each person among 14 people, factorization (i.e., adding five and seven before multiplying the total by 14) was more commonly used if the wording mentioned that the flowers were grouped in a bouquet than if it did not. Finally, Thevenot and Oakhill (2005, 2006) showed that the choice between two alternative solving algorithms is influenced by the cognitive costs of each strategy. Facing a problem statement where the solution was usually obtained by calculating the value of "$x - (y+z)$," they found that participants' preferences shifted in favor of the more economical sequential strategy "$(x-y) - z$" when presented with higher values.

Separately, these studies have all been accounted for within a given framework of arithmetic word problem solving; either the schema theory (Kintsch & Greeno, 1985), the situation problem model (Reusser, 1990; Staub & Reusser, 1995) or the semantic alignment framework (Bassok, 2001). However, taken together, these studies on wording effects, content effects, and re-representation processes display a range of findings that, to our knowledge, remain to be explained within a common model. To address this issue, we hereby propose a *Semantic Congruence* (SECO) model accounting for how the interactions between the solver's knowledge about the world (the world semantics) and the solver's knowledge about mathematics (the mathematical semantics) mediate the conceptual and procedural sides of arithmetic word problem solving. We believe that such a model should help pave the way toward the development of new instruction methods by providing a unified account of a range of effects whose considerable influence on students

CONTACT Hippolyte Gros hippolyte.gros@cri-paris.org Center for Research and Interdisciplinarity, Paris Descartes University, 12 rue de l'École de Médecine, 75006 Paris, France.

Supplemental data for this article can be accessed at publisher's website.

at all levels tends to be underestimated. Before further specifying the SECO model's inner workings, a description of the range of effects that current theories of arithmetic word problem solving do account for seems in order.

Arithmetic word problem solving theories

Numerous works have highlighted the fact that arithmetic word problems which can be solved using identical arithmetic operations may vary greatly in terms of solving difficulty, be they additive (Carpenter & Moser, 1982; Nesher, Greeno, & Riley, 1982; Riley, Greeno, & Heller, 1983) or multiplicative (Greer, 1992; Squire & Bryant, 2002; Vergnaud, 1983) problems. The two most prominent approaches of arithmetic word problem solving which have attempted to account for such effects are the schema and the situation model theories (see Thevenot & Barrouillet, 2015, for a review).

The schema model

The schema model (Kintsch & Greeno, 1985; Rumelhart, 1980; Schank, 1975; Schank & Abelson, 1977) posits that the resolution of arithmetic word problems relies on the creation, activation, and implementation of schemas. Schemas are defined as propositional data structures stored in long-term memory, as a result of repeated encounters with problems sharing the same structure. These operatory structures, once created, can be activated and implemented with numerical values from any given context (any cover story), thus providing the solver with a valid solving algorithm. According to this view, the solvers read the problem statement and "the verbal input is transformed into a conceptual representation of its meaning, a list of propositions" (Kintsch & Greeno, 1985, p. 111). The solvers then activate, in their long-term memory, the schema sharing the same propositional structure as the one in the problem statement. They then instantiate this schema with the specific numerical values of the cover story to interpret and solve the problem. For instance, in a compare problem, a sentence such as "Tom has three more marbles than Joe" cues a "have more than" propositional structure which uses three arguments: two corresponding to Tom and Joe's sets, and one corresponding to the quantitative proposition associated with the comparison (Kintsch & Greeno, 1985). According to Kintsch and Greeno (1985), this propositional structure can be implemented with the values of any problem using a "have more than" proposition and can be used to choose the solving algorithm.

However, the schema theory has been challenged by works showing that minor modifications within the wording of otherwise structurally identical problems led to significant differences in terms of solvers' performances. Notably, De Corte, Verschaffel, and De Win (1985) showed that modifying the wording of problems sharing the same schema impacted both their difficulty and the type of errors solvers make. For example, problems such as "Bob got 2 cookies. Now he has 5 cookies. How many cookies did Bob have in the beginning?" were only solved by 36% of the children in the study, whereas slightly reworded problems such as "Bob had some cookies. He got 2 more cookies. Now he has 5 cookies. How many cookies did Bob have in the beginning?" were solved by 55% of the children.

Another convincing piece of evidence showing the limitations of the schema model was brought by Thevenot and Oakhill (2005), who asked adults to solve problems such as "How many marbles do John and Tom altogether have more than Paul? John has 29 marbles, Tom has 13 marbles and Paul has 26 marbles." This problem is usually solved with the algorithm $(29 + 13) - 26 = 16$, which could be explained by the schema model by the fact that the word "altogether" activates a Combine schema $(29 + 13)$ and the words "have more than" activate a Comparison schema ($42 - 26$; Riley et al., 1983). However, the authors showed that when the numerical values were replaced by 3-digit numbers (e.g., replacing 29, 13, and 26 by 749, 323, and 746, respectively), participants tended to use another algorithm to solve the problem: $(749 - 746) + 323 = 326$. Indeed, since in both cases John has three more marbles than Paul, it would be easier to calculate the difference between John's and Paul's marbles and add it to the number of marbles Tom has. Yet, participants only used this strategy when the use of 3-digit numbers made it too difficult to calculate the solution using the other algorithm. This experiment suggests that participants were able to decide not to blindly apply the schemata activated by the problem and to construct an alternative problem representation instead.

Another argument showing the limitations of the schema theory came from Thevenot (2010), who asked participants to solve arithmetic problems and later presented them with an unexpected recognition task involving problems that were either identical to the source problems, inconsistent with the source problems, or that described the same situations using paraphrases. The results showed that paraphrastic problems had a higher recognition rate than inconsistent problems. Since, in paraphrastic problems, the propositional structure of the initial problems was lost by the paraphrasing, it follows that recognition was not solely based on a propositional representation, contrarily to what the schema view predicts.

Thus, additional interpretative processes are believed to come into play and modulate the solvers' performance. In this regard, effects of content—interpretative effects linked to the semantic content of the cover stories—have been shown to influence participants' performance in a way that is not accounted for by the schema theory (Coquin-Viennot & Moreau, 2003; De Corte et al., 1985; Gvozdic & Sander, 2017; Reusser, 1988; Vicente, Orrantia, & Verschaffel, 2007). This significant blindspot in the schema theory explains the need for a more comprehensive model accounting for the content effects reported in the literature.

The situation model approach

Due to these limitations, the schema theory has since lost ground against an alternative approach, which builds on the theoretical frameworks of mental models (Johnson-Laird, 1980, 1983) and situation models (Van Dijk & Kintsch, 1983). This approach originates from Reusser's model, the *Situation Problem Solver* (SPS), which applies the situation model approach to arithmetic word problem solving (1989, 1990, 1993; Staub & Reusser, 1995). The SPS model accounts for the integration by the solver of the set of information present in the problem statement. Namely, it proposes that reading a word problem results in the creation of an episodic situation model featuring every functional relation depicted within the text and presenting an analogous structure to that of the described situation (Reusser, 1990). For example, in Hudson's study (1983) mentioned in the introductory paragraph of this paper, the "How many more birds than worms are there" problem refers to a static episodic situation model where birds and worms are conceived of as two disjoint sets of entities, whereas the "How many birds won't get a worm" problem leads to the creation of a dynamic episodic situation model in which the relation between the two sets is highlighted (Staub & Reusser, 1995). The episodic situation model is then translated into a problem model containing the relevant structural elements and relations from the point of view of the question to be answered. This qualitative representation of the problem statement differs from the purely propositional structure proposed by the schema theory. According to Staub and Reusser (1995), this problem model is then reduced to its abstract mathematical gist, which can be translated into a solving algorithm.

Although it builds on the idea that solvers reason based on mental representations analogous to the situations described in the problem statements, the SPS model does not explicitly describe the processes that form those representations. Indeed, according to the situation model view, "the structure of a representation corresponds to the structure of what it represents" (p. 18244, Johnson-Laird, 2010). If a perfect structural correspondence is assumed between the representation itself and what is represented from the external world, this means that the former is presumed to be a faithful internalization of an external state. The processes through which this internalization is achieved are not explicitly in the scope of the SPS approach. In particular, the idea that background knowledge of an individual might influence the internalization process and eventually interfere with the faithfulness of the internalization relatively to the external situation is not a significant topic in the SPS model.

The notion that the structure of a representation is identical to the structure of what it represents is hardly compatible with the thought that one depicted situation could be interpreted differently by different individuals. In other words, saying that a problem statement is encoded as a representation whose structure is analogous to the problem statement's is tantamount to saying that only one representation can be encoded from a given problem, regardless of variations in interpretation that can occur over time or individuals.

The semantic alignment contribution

Other works have been more attentive to this issue, showing that solvers' prior knowledge strongly constrains the representations they construct, in an often detrimental way (Thevenot, 2017). Bassok et al. (1995) showed that the world knowledge regarding the entities involved in arithmetic problems influenced the transfer to isomorphic permutation problems; for instance, problems involving objects and people, such as caddies and golfers, spontaneously evoke an asymmetric structure ("get"), in which golfers are getting caddies and not the opposite since in our world, in most pragmatic contexts, people receive objects and not the other way around. In contrast, they showed that problems involving two sets of people (e.g., kids from two nurseries) evoke a symmetric structure ("pair"), in which children from both nurseries are paired together. These semantic relations between the problem elements thus constrain participants' representations of the problems.

Bassok, Chase, and Martin (1998) provided additional evidence for this claim, by giving participants the names of different types of objects and asking them to use these objects to create arithmetic word problems involving either an addition or a division. For objects linked by an asymmetric functional relation (e.g., a container/content relation between vases and tulips), participants created more division problems (e.g., the number of tulips divided by the number of vases) than additions. On the other hand, with objects belonging to the same superordinate category, such as tulips and daffodils, participants created mostly additive problems.

This issue is all the more important given how Bassok et al. (1998) showed that the association between subclasses of objects and specific solving strategies is reinforced throughout education by the exercises proposed in mathematical textbooks. They showed that a vast majority of division problems in math textbooks include elements linked by asymmetrical relations whereas additive problems feature elements belonging to categories of the same taxonomic level such as red and blue marbles. This reinforcement throughout the years of arithmetic school teaching may contribute to the development and strengthening of robust solving biases among learners, making it especially important to model these interpretative effects of content to better capitalize on them.

The semantic alignment framework (Bassok, 2001) aims at accounting for these interpretative effects of content. It goes beyond the SPS view by specifying how world knowledge regarding the entities involved in the problem influences its representation by the solvers. It proposes that the solvers' knowledge about the objects described in the problem cover stories leads them to abstract an *interpreted structure*. This structure varies from one problem statement to another, depending on the roles defined by the world knowledge regarding the entities, even when those roles are not relevant—or even deleterious—with regard to the mathematical structure of the problems and the task at hand. Thus, the structure that is abstracted from arithmetic problems can facilitate the resolution when the relations it entails are *semantically aligned* with the objective mathematical relations of the problem, that is when the problem's semantic structure can be used "to infer, by analogy, its objective mathematical structure" (Bassok, 2001, p. 402; Bassok et al., 1998).

For example, performing divisions on problems involving oranges and baskets will prove easier than performing divisions on problems involving oranges and apples, because division is semantically aligned with asymmetrical structures such as the one between containers (the baskets) and their content (the oranges). Supporting this view, Bassok, Pedigo, and Oskarsson (2008) showed that addition facts are activated when they are primed by categorically related words usually associated with addition (e.g., the tulip–daisy pair is semantically aligned with addition), but not in cases of misalignment, when they are primed by unrelated words and are misaligned with addition (e.g., hens and radios are not usually connected in an addition model). This was confirmed in an ERP study by Guthormsen et al. (2016) who showed N400 and P600 effects indicating a disruption of conceptual integration when participants were presented with misaligned problems (e.g., a problem in which flowers and vases were added together). These results indicate that, in case of semantic alignment, the semantic content of a problem statement can provide crucial clues to the solvers.

Alternative encodings and re-representation

The strengths of the previous approaches are their versatility and their ability to each account for a range of effects documented in the literature. However, it seems that one crucial question remains open: how is it possible to solve a problem whose semantic content is misaligned with its solution? How can one ignore those misleading clues and go beyond their initial encoding of a problem statement to reach a solution? Overcoming semantic misalignment would mean abstracting a new, different structure of the depicted situation. For instance, encoding a problem with caddies and golfers as a distributive structure where golfers are assigned to caddies instead of the opposite.

However, the issue of whether several alternative interpreted structures can be encoded from the same problem statement, by different individuals or by one individual over time, has yet to be considered. Ross and Bradshaw (1994) showed that the initial interpretation of an ambiguous story could be influenced by the beforehand presentation of another story sharing some degree of similarity with the latter. This suggests that two different semantic structures can be abstracted from a same situation, depending on participants' past experiences and prior knowledge.

Furthermore, studies on re-representation showed that it is possible for the solvers to turn their initial representation into a new one, allowing them to overcome their initial inappropriate interpretation and find the solution (Davidson & Sternberg, 2003; Gamo, Sander, & Richard, 2010; Sander & Richard, 2005; Vicente et al., 2007). For example, to facilitate the solving of a change problem in which a quantity is added or subtracted from an unknown start set, solvers can represent the problem in terms of a part-whole structure and turn it into a search for the unknown part (Riley et al., 1983). Thus, it is important to tackle what precisely happens when a solver's initial encoding of a problem statement fails to trigger the use of an appropriate solving algorithm, and to get a better understanding of how solvers might overcome an earlier inadequate representation and recode the same problem. Bearing this issue in mind, we wish to build on the SPS model and on the notion of interpreted structure in order to provide a unified model addressing the processes involved in arithmetic problem solving.

The Semantic Congruence (SECO) model

The SECO model is based on the notion of semantic congruence in arithmetic word problem solving, which it defines and operationalizes by accounting for the interactions between world semantics, mathematical semantics, and algorithms. Within the SECO model (Figure 1), the product of the interaction between world semantics and mathematical semantics needs to be put in correspondence with an algorithm, by means of an interpreted structure.

Components

The components depicted in the SECO model are characterized below; they will be further exemplified in a second phase through six case studies.

- **Problem statement**: The problem statement is a text describing the elements of the problem and the situation(s) in which they interact as well as their relations and associated values.
- **World semantics**: World semantics is characterized by the solver's non-mathematical, daily-life knowledge about the elements of the problem statement as well as the relations between them. For example, world semantics may include knowledge that flowers can be put into vases, that there is a co-hyponym relation between oranges and apples, or that to go from the first to the third floor of a building one must pass by the second floor first. There is indeed a broad literature showing that understanding, reasoning, decision-making and problem solving are influenced by an individual's knowledge regarding the entities involved and their relations (e.g., Bassok, 2001; Carey, 2009; Gelman, 2003; Gentner, 1988; Goswami & Brown, 1990; Johnson-Laird, 1983; Kotovsky, Hayes, & Simon, 1985; Stanovich, 1999; Van Dijk & Kintsch, 1983).
- **Mathematical semantics**: Mathematical semantics is characterized by the solver's mathematical knowledge that is applicable to the problem statement. For example, mathematical semantics may include knowledge that to calculate the size of a set, one needs to add the size of all its subsets, or that to evenly share a collection of objects among several sub-collections, one needs to divide the

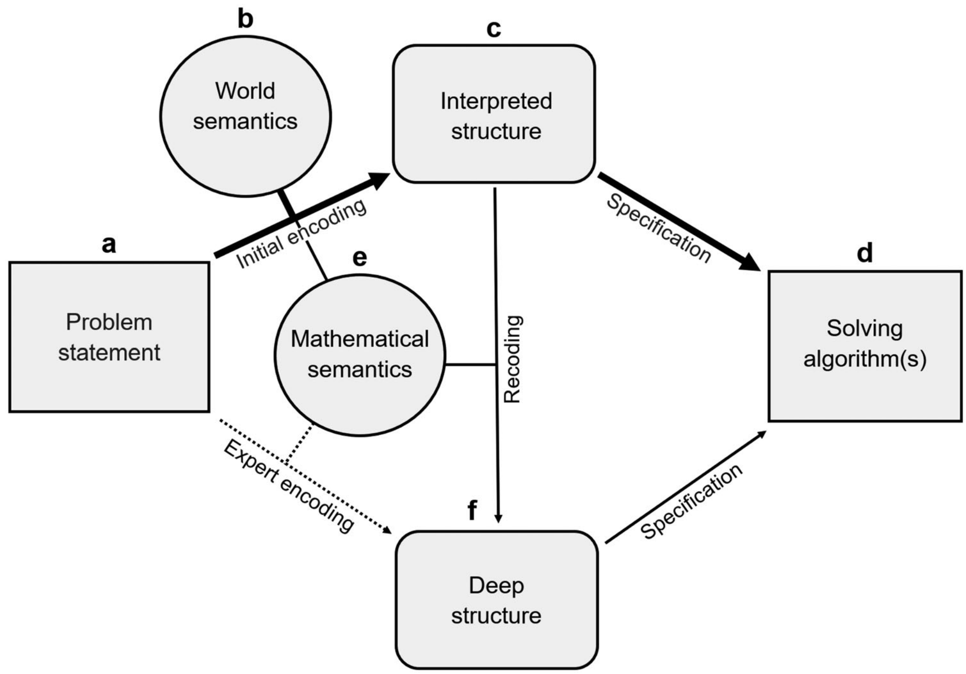

Figure 1. The SECO model.

number of elements in the collection by the number of sub-collections.

- **Interpreted structure**: The interpreted structure is abstracted from the problem statement, integrating pieces of information present in the text with the properties, relations, and constraints inferred from the world semantics. This notion stems from Bassok and colleagues' research (Bassok & Olseth, 1995; Bassok et al., 1995, see above). Since the mathematical semantics evoked by the problem statement is activated during the encoding, the interpreted structure can feature algebraic values or be instantiated by the numerical values. For example, world semantics about fruits will lead co-hyponyms such as oranges and apples to be encoded as subsets of a superset of fruits.
- **Solving algorithm(s)**: A solving algorithm is a finite, unambiguous set of actions that leads to the correct answer when properly executed. Multiple solving algorithms may stem from a given problem statement (e.g. De Corte et al., 1985; Gamo et al., 2010; Große & Renkl, 2006; Kouba, 1989; Leikin & Lev, 2007; Thevenot & Oakhill, 2005).
- **Deep structure**: This notion stems from Chi and colleagues' work (Chi, Feltovich, & Glaser, 1981). We define it as the semantic structure integrating the elements of the problem that are relevant for its resolution and describing their relations. This structure does not rely on world semantics but on mathematical semantics. It has been designated as "the objective mathematical structure" (Bassok, 2001), or as "the principle of the problem" (Ross, 1987); for non-mathematical problems, the corresponding notion is "the problem space" of an expert solver (Newell & Simon, 1972).

Processes

The processes depicted in the SECO model are characterized as follows:

- **Initial encoding**: This process describes how the problem statement is abstracted into an interpreted structure depending on the world and mathematical semantics evoked by its wording. The world semantics activated by the problem statement constrains the representation of the depicted situation, either by highlighting or by overshadowing specific relations between the problem's entities. Similarly, the mathematical semantics evoked by the problem statement also shapes the mathematical relations represented in the interpreted structure.
- **Specification**: This process describes how an interpreted structure may be specified into an algorithm, as a result of the relations it describes and the numerical values it features. When the relations depicted in the interpreted structure hold a mathematical meaning, they can be translated into relevant operations through this specification process. Not every interpreted structure can be specified into a relevant solving algorithm, since the relations highlighted during the encoding process may not be relevant, and the encoded values may not be the ones needed to solve the problem. A deep structure, on the other hand, may be specified into any relevant algorithm, since it depicts every relevant relation, independently from the influence of world semantics, contrarily to an interpreted structure.
- **Expert encoding**: The expert encoding describes the hypothetical process that may happen when experts initially encode problems within their domain of expertise. As stressed by Chi et al. (1981) experts are believed to be

able to disregard the cover story of a problem and directly encode its deep structure. According to this view, an expert may use mathematical semantics and disregard world semantics to directly abstract the deep structure from the problem statement.

- **Recoding**: Since not every deep structure can be specified into a relevant solving algorithm, the recoding describes how, when the initially encoded interpreted structure cannot be translated into an appropriate, tractable algorithm, a new representation can be abstracted by recoding the interpreted structure. This process is akin to the re-representation said to be necessary to overcome difficulties in arithmetic problem solving (Vicente et al., 2007). It relies on mathematical semantics to recode the situation and build a new structure, closer to the deep structure of the problem. It is a costly process that does not systematically happen.

Inner workings

The SECO model integrates these notions in the following way: it posits that when reading a problem statement (a), the lay solvers will initially encode the problem according to the world semantics (b) as well as to the mathematical semantics (e) evoked by the problem statement, from which they will abstract an interpreted structure (c). This interpreted structure is therefore semantically aligned with the solvers' knowledge about the elements present in the problem statement and can differ from one solver to another for the same problem statement, depending on the state of their world and mathematical semantics. Because it holds a mathematical meaning, this interpreted structure may be specified into an algorithm (d). This algorithm stems from the procedural knowledge that is attached to the mathematical semantics activated by the problem statement. In cases in which no tractable algorithm can be derived from the interpreted structure encoded, the solver faces a dead-end and the need for a recoding process arises. Such a process would appeal to mathematical semantics (e) and not to world semantics, in order to encode a new representation consistent with the deep structure (f) of the problem and thus allow the use of a new algorithm as a result. Contrarily to the interpreted structure from which no tractable algorithm might be derived, this deep structure can be specified into any relevant solving algorithm (d). Finally, the model also introduces the possibility that an individual with sufficient expertise regarding a specific type of problem might directly abstract a deep structure (f) from a problem statement (a), without first extracting an interpreted structure (c) influenced by world semantics (b).

SECO underlines a key aspect of arithmetic word problem solving consisting in the congruence between the semantic knowledge evoked by the problem statement and the mathematical semantics required to find its solving algorithm. If the world semantics attached to the elements in a problem statement is not congruent with the mathematical semantics required to solve the problem, the initial interpreted structure will not be translated into a valid solving algorithm. Indeed, only the mathematical semantics congruent with the world semantics evoked by the problem statement will be used during the initial encoding of the problem. In cases where the relevant mathematical semantics is not congruent with the world semantics evoked, an extra recoding step is necessary to recode the interpreted structure into a new representation closer to the deep structure of the problem, making the process longer and more difficult.

As in the SPS model, SECO considers that a mental representation of the situation is abstracted when reading an arithmetic word problem. However, contrarily to this model, SECO does not consider that this representation maps onto the structure of the world: by integrating the role of world and mathematical semantics in the encoding of the problem statement, SECO accounts for the fact that there is no unique way to mentally model a problem statement. A situation can be encoded differently by different individuals, and the abstracted structure may be recoded into a new representation if need be.

Accounting for existing results: Case studies

In order to better understand SECO's contribution in contrast to the current models of arithmetic problem solving, we propose to tackle representative results in the field through SECO's lens and compare it to the accounts of these results by the two most prominent models of arithmetic word problem solving, the Schema model and the Situation Problem Solver model. As our presentation of SECO shows, its main contribution resides in its depiction of the influence of world semantics on solving strategy choice as well as of the necessity to semantically recode the problems in case of failure. While SECO does not intend to resort solely to world semantics to account for every possible variation in arithmetic problem solving, as other sources of differences exist (e.g., algorithm computation abilities or reading comprehension), its central added value consists in its depiction of the influence of world semantics on the encoding, recoding and solving of the problems.

We now assess SECO's unique ability to explain the effects reported by a set of six studies mentioned in the introductory section of this paper and presenting representative results in the field. We believe that altogether, these studies prove challenging to the existing models of arithmetic problem solving. We first present two cases illustrating the key issue of the influence of world semantics on the selection of a solving strategy. The following two case studies then showcase the other central feature of the SECO model: its depiction of the existence of a recoding process for semantically incongruent representations. The last two case studies show how SECO proposes a new take on classical rewording effects, from which important educational implications arise.

World semantics issues

The first two studies we detail illustrate the key influence one's knowledge about the world can have on one's problem solving performance. They describe examples of the effect the content of a problem statement can have on the interpretative processes at play. In other words, they showcase the role of world semantics in arithmetic word problem solving.

Case 1: Bassok et al.'s account of interpreted structures

Empirical findings and authors' perspective. Compelling evidence of the influence of world semantics on the interpreted structure have been provided by Bassok et al. (1995). Participants who were unable to solve an initial permutation problem were presented with a short lesson accompanied by a training problem and its solving equation: $\frac{1}{n(n-1)(n-2)}$ (n being the size of the set of elements being assigned). The participants then had to solve a transfer problem using the same algorithm. The main result was that participants who were first trained on a problem involving an assignment of objects to people (O→P) had a dramatically higher success rate when they transferred the solution to other O→P problems (89% of success) than those who had to transfer the solution to "people assigned to objects" (P→O) problems (0% of success).

According to the authors, the participants interpreted the structure of the problems by using their world knowledge about the roles of the entities involved in the problems, i.e., they spontaneously interpreted the problem as a situation in which "objects are given to people" and constructed different interpreted structures depending on which entities were described. The semantic (mis)alignment between the training and transfer problems' interpreted structures accounted for the participants' high (or low) success rate in the transfer problems.

SECO's account of the results. Because it details how an interpreted structure is encoded according to the world and mathematical semantics, the SECO model can account for this result, see Figure S1 in Supplemental Materials (transfer to objects-to-people problems) and Figure S2 (transfer to people-to-objects problems). As situations where objects are assigned to people are much more common in daily-life than situations where people are assigned to objects, in SECO, the world semantics (b) regarding the assignment of elements fosters the idea that objects are usually assigned to people. Therefore, when reading the problem statement (a), the world semantics (b) should, in both "objects to people" and "people to objects" problems, result in an interpreted structure (c) in which objects are assigned to people.

This interpreted structure leads the participants to implement the algorithm $\frac{1}{n(n-1)(n-2)}$ (d) with the value corresponding to the size of the set of inanimate objects whereas they should be thinking in terms of which set is being assigned to the other. Indeed, given that participants received limited training, it might be that they did not really understand the solving procedure in the training problem, and thus their mathematical semantics (e) regarding the assignment did not comprise the mathematical notion of "draw without replacement within a set." Instead, they simply implemented the training algorithm by mapping the semantic roles of the training and transfer problems, and only considered the fact that n was the size of the set of assigned objects in the first problem. Thus, they transfer the algorithm they learned by replacing the n value by the number of inanimate objects, even if the set of people is the one being assigned to the set of objects. This leads to correct use of the algorithm in "object to people" transfer problems (Supplementary Figure S1) but not in "people to objects" transfer problems (Supplementary Figure S2) and accounts for the dramatic contrast between the transfer rates in both conditions (0% vs. 89%).

Case 2: Coquin-Viennot and Moreau's account of semantic constraints

Empirical findings and authors' perspective. In their study bearing on the use of factorization and expansion algorithms among third and fifth graders, Coquin-Viennot and Moreau (2003) showed that problems such as "For a prize-giving, the florist prepares for each of the 14 candidates 5 roses and 7 tulips. How many flowers does the florist use in total?" were less often solved using factorization (44% among fifth graders) than problems identical in every aspect except for the presence of a superordinate structuring term such as "a bouquet": "For a prize-giving, the florist prepares for each of the 14 candidates a bouquet made up of 5 roses and 7 tulips." (68% among fifth graders).

This study illustrates how slight modifications in the wording of isomorphic problems can influence the initial encoding. The interpretation proposed by the authors was that the presence of the term "bouquet" favored participants' perception of the two subsets as parts of the same superset and led them to combine the sets into a single entity. We propose a complementary and more systematic explanation using the SECO architecture.

SECO's account of the results. In SECO's view, the use of the word "bouquet" in the problem statement evokes the world semantics stating that a bouquet is a group of flowers, which is compatible with Coquin-Viennot and Moreau (2003) interpretation. The SECO model would account for these results as depicted in Supplementary Figure S3 (problem statement without the "bouquet" term) and Figure S4 (problem statement with a structuring term).

Since the "no bouquet" problem statement (Supplementary Figure S3, a) mentions roses and tulips, the world semantics (b) regarding those elements (i.e., "roses and tulips are two different kinds of flowers") is activated and favors the encoding of roses and tulips as two disjoint sets in the interpreted structure (c), making the grouping of roses and tulips together less salient. The abstracted interpreted structure (c) thus leads most of the participants to use the expansion algorithm (d) congruent with the representation of tulips and roses as two distinct sets "$(14 \times 5) + (14 \times 7)$."

By contrast, in order to use the factorization algorithm "$14 \times (5 + 7)$," a solver is either required to infer that tulips and roses can be grouped together (e.g., in a bouquet constituted of different flowers), despite the absence of any structuring cue, or to recode the situation (c) according to mathematical semantics (e**) stating that a superset consisting of m sets of x elements and m subsets of y elements has the same size as a superset consisting of m subsets of "$x + y$" elements, so as to abstract a deep structure (f) of the problem. This deep structure highlights the two different grouping strategies (grouping by individuals or grouping by types of flowers) and is thus congruent with both the factorization algorithm and the expansion algorithm.

On the other hand, the resolution of the problem mentioning a structuring element (the bouquet) leads to different steps as detailed in Supplementary Figure S4. When the problem statement (a) mentions that the tulips and the roses are grouped together and form a bouquet, then the world semantics (b) related to the bouquet can also be used, in addition to the world semantics related to roses and tulips as flower species. Referring to a bouquet activates the notion of grouping within a single set and helps the solver encode an interpreted structure (c) increasing the saliency of the two subsets of flowers as parts of the same "bouquet" set, compared to when the structuring element was not mentioned in the wording. The interpreted structure (c) leads the solver to calculate the total number of flowers by adding the number of flowers in each bouquet. Therefore, the factorization strategy $14 \times (5 + 7)$ is the one being mostly used by solvers in this situation. The use of the expansion algorithm "$(14 \times 5) + (14 \times 7)$" is less frequent on such problems and can be the consequence of participants focusing on the distinction between the two types of flowers, roses and tulips, that leads them to count those separately instead of counting the number of flowers within one bouquet first. Alternatively, it can also be the consequence of their explicit use of mathematical semantics (e**) regarding expansion and development.

Thus, within the SECO model, the difference between the two problem statements results in a difference between the world semantics evoked by the statements during the encoding process. Different world semantics result in different constraints influencing the encoding of the problems, which lead to different interpreted structures being abstracted, each of them congruent with a specific solving algorithm (expansion when no structuring element is present in the problem statement and factorization otherwise).

Other models' accounts and their limitations regarding cases 1 and 2

The influence of world semantics displayed by these two case studies is an effect that clearly falls outside the scope of the schema model. In Coquin et al.'s case, there is no theoretically based reason justifying that the addition of the term "bouquet" could influence the selection of a completely different problem schema. Similarly, in Bassok et al. (1995) case, the problem schema in the "objects to people" situation should be the same as the one in the "people to object" version, since the only change introduced between the two problems was the semantic nature of the entities constituting the two sets (either people or inanimate objects).

In the original schema model, the reader extracts the numerical values and their relations by focusing on the propositional structure of the text (Kintsch & Greeno, 1985). In Bassok et al.'s work (1995), the sentences "The president randomly assigns students to prizes" and "The president randomly assigns prizes to students" have the same propositional structure and should have activated the same schema. However, because one sentence implied that objects were assigned to people, and the other that people were assigned to objects, participants' strategies differed between the two problems. The schema theory alone cannot account for this performance difference without being updated to take into account solvers' knowledge about the problems' entities.

Similarly, the SPS model does not directly integrate the idea that one's knowledge about the entities in a problem could influence the episodic situation model constructed to solve it. Instead, it postulates that the episodic situation model that is built depends on the presentational structure of the problem (text order, narrative point of view, presence of an explicit question, explicitness of relevant relations and so forth) but not on the general knowledge imbued in the problem (Staub & Reusser, 1995). The SPS model relies on the Situational Model assumption that the structure of a representation maps onto the structure of what it represents (Johnson-Laird, 2010), therefore suggesting that there is only one episodic situation model for each problem, regardless of participants' previous knowledge about the entities featured in a problem. The SECO model, on the other hand, provides a satisfactory account of those results by proposing that the world semantics evoked by a problem also depends on the semantic nature of the elements featured in the problem statement.

Recoding issues

While the first two case studies focused on the mechanisms at play during the initial encoding of a problem statement and their consequences on the solving performances, the next two case studies highlight how an interpreted structure resulting in a dead-end can be recoded in certain conditions. In other words, they focus on participants' relative ability to change their initial representation in situations in which multiple mathematical encodings of the same problem statement are possible: different, equally valid representations emphasizing distinct relations.

Case 3: Thevenot and Oakhill's account of alternative representations

Empirical findings and authors' perspective. Studying the influence of number size on the use of solving algorithms, Thevenot and Oakhill (2005) shed light on the factors triggering the recoding of an interpreted structure into a new representation. They investigated the strategies used to solve *compare* problems by using an operand-recognition paradigm consisting in interrupting the presentation of the problem statements to ask participants whether they recognized specific numbers. Recognition performance was used to determine if these numbers were currently maintained in working memory or if they had already been used in calculation and had thus started to fade from memory.

They used problems such as "How many marbles does John have more than Tom and Paul together? John has x marbles, Tom has y marbles and Paul has z marbles." The authors' findings show that participants used the grouping algorithm "$x-(y+z)$" when the task was not especially demanding due to the problem's values being small, whereas they used the more economical sequential strategy "$(x-y)-z$" when the use of larger values implied that the task had higher cognitive costs. Indeed, the second strategy is less cognitively demanding because performing two successive subtractions allows the solvers to complete a subgoal "$x-y$" while reading the problem, and the result of this operation can be maintained in working memory during the rest of the problem instead of the two initial values (x and y).

On the other hand, calculating the value of "$y+z$" and then subtracting it from the value of x requires to maintain the value of x in working memory until the end of the text and the final operation. In other words, if the values 636, 345 and 123 appear in that order in the problem statement, then it is easier to first calculate the value of "(636 − 345)" while reading the text and then subtracting 123 from the result later on than to memorize the three values to calculate "636 − (345 + 123)" at the end of the problem statement.

SECO's account of the results. Within the SECO model (see Supplementary Figure S5), this effect follows from the fact that the problem statement (a) mentions marbles that are grouped together and then compared. The interpreted structure (c) thus features two disjoint sets: one corresponding to John's marbles, and the other one to Tom and Paul's put together. This interpreted structure (c) is semantically congruent with the grouping algorithm "$x-(y+z)$" (d) that is preferentially used for problems with small values.

When computing the algorithm becomes impossible because of the larger x, y and z values, some participants need to recode the situation to avoid maintaining the three values in memory. By focusing on the mathematical knowledge regarding parentheses removal (e**), according to which "$x-(y+z)$ is equivalent to $x-y-z$," participants can recode their initial representation into an alternative representation closer to the deep structure (f) of the problem, in which Tom and Paul's sets are perceived as two independent sets that can successively be removed from John's set. They can then switch to the more economical sequential algorithm "$(x-y)-z$" (d). In other words, difficulty to compute the algorithm triggered a re-elaboration process that focused on the mathematical semantics to recode the problem's representation.

Case 4: Gamo et al.'s account of world semantics constraints and semantic recoding

Empirical findings and authors' perspective. In addition to being another illustration of the central role of world semantics on arithmetic word problem solving, the study that Gamo et al. (2010) conducted in fourth and fifth grade classrooms provides valuable insight into the semantic recoding of the initial, inadequate representation of a problem into a new, more polyvalent one. In their study, Gamo et al. used problems that all shared the same formal deep structure, but that involved different types of elements. When the elements were known by the solvers to be unordered entities, such as marbles, scissors or pens, the authors predicted that the participants would abstract an interpreted structure emphasizing the cardinality of the situation, such as an embedded sets structure. This structure was shown to lead the participants to use a 3-step algorithm to solve the problems. For example, the problem "John bought an 8-Euro exercise book and 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?" was preferentially solved using the 3-step algorithm consisting in calculating the price of the pen and the price of the scissors before adding them up: 14 − 8 = 6; 8 − 3 = 5; 6 + 5 = 11.

On the other hand, when problems involved ordered units, as is the case in problems involving age, where events are ontologically ordered on the line of time, the authors predicted that the participants would abstract an interpreted structure emphasizing the ordinality of the situation, such as a timeline where different events are represented as positions on an axis. This axis-based interpreted structure would make it possible for the participants to use a different solving algorithm. For example, the problem "Antoine took painting courses at the art school for 8 years and stopped when he was 14 years old. Jean began at the same age as Antoine and took the course for 3 fewer years. At what age did Jean stop?" was predominantly solved using a shorter, more efficient 1-step algorithm: 14 − 3 = 11.

Indeed, the fact that the problem involves durations makes it easy to see that since Jean and Antoine started taking the course at the same age, then the difference between the number of years they each followed the course is equal to the difference between the age at which they stopped taking the course. Thus, the problem can be solved without calculating their age when they started taking the class. Both problems could be solved using both algorithms, but depending on the elements featured in the problems, participants preferentially used one or the other of the two strategies.

In the first experiment of the study, the authors studied the conditions allowing for strategy change. They divided the participants into two groups, both of which had to complete a pretest and a post-test in which they had to solve similar problems using only one arithmetic operation. Between the two tests, one of the groups followed two 60-min training sessions during which the children were instructed to compare the two strategies and incited to see how the 1-step algorithm could be used even on problems with unordered entities. They were explicitly trained to identify their initial semantic representation and they were shown a visual representation of the deep structure of the problems to help them recode their initial encoding of the situation. The other group did not receive such training.

The two main findings were that children did solve problems differently depending on the world semantics they evoked, and that teaching the children to use both strategies by focusing on the mathematical relations between the entities described and by studying the deep structure of the problems yielded significant result in increasing their ability to use the shorter 1-step algorithm on problems with unordered elements.

SECO's account of the results. These findings are a perfect fit within SECO's framework, since they show both how mathematical and world semantics interact in the encoding of the problem statements into an interpreted structure, and how this interpreted structure then either leads to the use of a semantically congruent solving algorithm or is recoded to allow the use of a semantically incongruent solving strategy. Indeed, in the case of an age problem (see Supplementary Figure S6), the world knowledge (b) relating to how time events are usually conceptualized (as transitions between positions on a timeline) is evoked by the problem statement mentioning ages (a). This leads the children to encode an interpreted structure (c) in which the events described are represented along a timeline, which lets them directly compare the ages at which they each stopped attending the classes. This structure can then be specified into the 1-step algorithm (d) congruent with it.

On the other hand, when reading a problem with unordered elements (see Supplementary Figure S7), Gamo et al. (2010) indicate that the encoding is influenced by the students' knowledge (b) that the elements can be grouped together in any order, and that, for example, the scissors can be indifferently grouped with the pen or with the notebook. The resulting interpreted structure (c) has an embedded set structure that leads the students to calculate the value of each subset (the price of each item). This structure can then only be specified into the 3-step solving strategy (d). In order to use the shorter 1-step strategy, the students needed to use mathematical semantics (e) and recode their representation into a new, more polyvalent one (f). This explains why the only group who increased their performance in using the 1-step algorithm on problems with unordered entities was the one that followed a training based on the mathematical principle behind the use of the 1-step algorithm and the study of the deep structure.

Other models' account and their limitations regarding cases 3 and 4

These two last case studies showed that when the initial encoding of a problem statement does not lead to a satisfactory solving algorithm, a recoding may happen to encode a new representation congruent with a valid algorithm. As mentioned previously, in Thevenot and Oakhill's case, the idea that a problem could be solved differently depending on whether it features low or high values falls beyond the scope of the schema theory. Indeed, if a schema is constructed from the text-base, then two text-bases differing only by the range of their numerical values should result in two identical schemas being used. Even though it could be argued that students are switching from a schema to another depending on the values provided in the problem statement, such a claim would require a theoretical extension of the schema model accounting for the conditions under which such a switch can occur.

Similarly, if the SPS model predicts that one constructs a representation whose structure is that of the described situation, then why would two different representations be constructed based on the same situation? None of the aforementioned models of arithmetic word problem solving directly predicts that an encoding can be recoded depending on how efficient the algorithm it leads to is.

Finally, regarding Gamo et al.'s results, the schema theory may state that some problems correspond to a schema (the so-called ordinal problems) and some do not (the so-called cardinal problems). However, because this theory does not take the structure of the solver's representation into account, it provides no basis to explain why such a schema would only be used on some problem statements and not on others. Specifically, without these semantic features, there is no a priori reason to predict that words such as "age," "during" or "years" would activate a schema corresponding to the 1-step algorithm whereas words such as "scissors," "pen" or "book" would fail to do so.

On the other hand, the situation model approach states that a representation analogous to that of the situation described is constructed and used as a basis for reasoning. Because of that, this theory can explain why different problems can be represented differently and thus lead to the use of different algorithms, but the SPS model does not refer to the fact that solvers interpret the situations through the lens of their own previous knowledge. In other words, the situation problem view does not model the constraints imposed by world semantics on the encoding of arithmetic word problems.

Interestingly, it can be noted that the influence of general semantic dimensions such as the cardinal versus ordinal distinction is compatible with the semantic alignment framework. However, in the semantic alignment framework, the question of the recoding of semantically incongruent representations has not been addressed, and SECO's predictions regarding the students' ability to perform a semantic recoding when given appropriate guidance fall beyond this framework. Thus, the fact that the participants were able to solve the problems with unordered entities using the 1-step algorithm after the training sessions is not predicted by the semantic alignment framework, whereas SECO's take on semantic recoding aided by mathematical semantics offers a reasonable explanation of the effect.

Rewording issues

Several works have highlighted how small modifications in the wording of structurally isomorphic problems could result in significant performance disparities (Cummins, 1991; Cummins, Kintsch, Reusser, & Weimer, 1988; Davis-Dorsey, Ross, & Morrison, 1991; Staub & Reusser, 1992; Stern & Lehrndorfer, 1992; Vicente et al., 2007). Such effects have considerable educational implications since they illustrate how minor phrasing variations can drastically help (or hinder) the students' understanding of a given problem. As such, they constitute a promising route to assist students in overcoming some of the obstacles they meet in arithmetic word problem solving. Here, we focus on two studies showcasing such rewording effects, to illustrate how SECO can also account for such emblematic results by depicting the changes they entail in the interpreted structures abstracted.

Case 5: Hudson's account of children's understanding of differences between sets

Empirical findings and author's perspective. In his seminal work on numerical differences, Hudson (1983) compared two formulations of comparison problems that led to considerably different levels of performance. Kindergarten children were told there was, for example, "5 birds and 3 worms," and they were asked either "How many more birds than worms are there?" (25% of correct answers among kindergarteners) or "How many birds won't get a worm?" (96% of correct answers among kindergarteners). The author explains that the use of "won't get" reduced the misinterpretation of the "how many more than" construction by highlighting the one-to-one correspondence between the given sets.

SECO's account of the results. The SECO model accounts for those results in the following way. As depicted in Supplementary Figure S8, an interpretation of Hudson's findings within the model would be that the sentence "how many more birds than worms are there" in the problem statement (a) evokes aspects of world semantics (b) emphasizing the difference between the two sets of elements (knowledge that birds and worms are two different animal species) thus inducing a comparison between the two groups of elements, without specifying how these two groups should be compared. In contrast, as depicted in Supplementary Figure S9, the wording of the problem statement (a) in the "won't get" condition emphasizes the pairing relation between birds and worms and evokes a different aspect of world semantics (b) (i.e., "birds eat worms") which promotes the mapping between the two sets within the interpreted structure (c). Thus, in the "more" condition, the interpreted structure (c) consists in two disjoint sets of elements and provides no hint that would trigger a subtraction algorithm.

By contrast, the interpreted structure (c) in the "won't get" condition affords a one-to-one mapping between 3 birds and 3 worms. The "won't get" condition, therefore, evokes an interpreted structure that is semantically congruent with an efficient strategy (d), namely counting from 3 to 5. In the "more" condition, recoding the interpreted structure into a deep structure (f) of the problem remains possible, but requires using mathematical semantics (e**) about subtraction, which is not systematically acquired at this early age, thus explaining the low performance on this task (25% among kindergarteners). While Hudson accounted for this finding by stating that comparable constructions of the general form "how many more […] than?" tended to be misinterpreted, SECO provides an account of this effect in terms of representational differences.

Case 6: De Corte et al. (1985) account of rewording effects

Empirical findings and authors' perspective. De Corte et al. (1985) used combine, compare and change problems to study the effects of conceptual rewording on first and second graders' performance, and brought further evidence of the positive effects of specific forms of rewording. For each problem, they compared a "standard" version with a "reworded" version that stated more explicitly the relations between the sets to make them clearer for young students. For example, one of the compare problems they created was "Pete has 8 apples. Ann has 3 apples. How many apples does Pete have more than Ann?". They compared students' performance on this problem and on its reworded version: "There are 8 riders but there are only 3 horses. How many riders won't get a horse?".

Results showed that 47% of first-graders managed to solve the compare problems in their standard version, whereas 70% of them managed to solve the reworded version. With a rate of success of, respectively, 76% on standard compare problems and 90% on reworded compare problems, second-graders also benefitted from the conceptual rewording, although to a lesser extent. The authors explained this difference between the two conditions by stating that only the "won't get" condition provided enough linguistic cues to compute the difference between the sets, whereas the "more" condition remained ambiguous to inexperienced solvers.

SECO's account of the results. SECO provides a complementary account of these results. In the standard version (Supplementary Figure S10), the problem statement (a) does not evoke any aspect of world semantics that could help with the matching of the two sets in the interpreted structure (c). Thus, students who have not sufficiently acquired the mathematical semantics (e**) regarding the calculation of the difference between two sets will fail to solve the problem. This explains why standard compare problems had a low rate of success for first-graders and a higher one for second-graders.

On the other hand, the reworded problem statement (Supplementary Figure S11) evokes knowledge about riders and horses (b) namely the information that a rider is supposed to ride a horse. The interpreted structure (c) thus features the pairing of the three horses with their respective riders and makes it easier to understand how to count the horseless riders remaining. The mathematical semantics (e**) is not necessary in this case to solve the problem, which explains why the performance rate was higher in both age groups.

Other models' account and their limitations regarding cases 5 and 6

As stated by Vicente et al. (2007), the computational models using problem schema as a basis to explain word problem solving behaviors have struggled to systematically explain the rewording effects of studies such as the two presented above, due to the relatively weak elaboration of the first text-processing stage in their models. Both in Hudson (1983) and in De Corte et al. (1985) study, the initial problems and their reworded counterparts shared the same structure according to Riley et al. (1983) classification of additive one-step problems. However, small modifications in the wording of the problem statements resulted in significant performance disparities, an effect that the schema model would struggle to account for.

On the other hand, the SPS model focuses on the idea that a representation, the episodic situation model, is built featuring the relations depicted in the problem statement. According to Staub and Reusser (1995), this representation is different in the two conditions, since the "won't get" situation imbues the difference with real-world meaning, whereas the "more" condition only refers to a static, abstract situation. This suggests that the SPS could have predicted such rewording effects relying on an elaboration of the semantic relations described in the text, since it made the relations between the sets more salient, which explains why the representation was more accurate and led to a higher success rate in the "won't get" condition. However, it can be noted that any rewording effect capitalizing on prior knowledge, such as replacing computers and secretaries by two sets of doctors in Bassok et al. (1995) work, would fall beyond the scope of the SPS model.

Conclusion

When taken together, these six case studies show how SECO can account for varied results within a unified model. While explanations for these results have been provided by one of the already existing theories of arithmetic word problem solving, it appears that none of the aforementioned models can account for all of them simultaneously. In our view, one of SECO's strengths is that it provides an original integrative framework for the existing results in the literature.

SECO's added value

The current paper proposes a model detailing the processes at play in arithmetic word problem solving and accounting for how algorithms are found by solvers and how their performances may differ depending on the task. SECO describes how a problem statement is encoded into an interpreted structure according to the world semantics and the mathematical semantics, and how this structure can either be specified into an algorithm when congruent with one, or recoded into a deep structure thanks to mathematical semantics in order to solve a semantically incongruent problem. We illustrated its ability to explain a wide range of effects by confronting SECO, post hoc, to previous studies presenting challenging results that had yet to be accounted for within a unified framework.

The idea that there exist different possible encodings of a situation described in a specific problem is central in the SECO model, yet this view appeared only recently in the literature. Ever since Riley et al. (1983) work and their taxonomy of additive word problems, the view that a word problem can be reduced to its objective mathematical structure and that two isomorphs of the same problem can thus be considered as equivalent in terms of difficulty for the solvers was abandoned in favor of an approach putting more emphasis on the way different arithmetic word problems are interpreted. It has for example been highlighted by Riley et al. that *combine* and *compare* problems can be approached very differently by the solvers, even when both are subtraction problems involving the same numerical values.

However, in the Riley et al. (1983) view, each situation is attached to only one category in a taxonomy encompassing all problems, therefore suggesting that there is only one way to interpret a given situation. Similarly to Socrates' depiction of the human ability to "separate things according to their natural divisions, without breaking any of the parts the way a clumsy butcher does" (Plato, trans. 2009, p. 64), this view presumes that there exists a natural breakdown of the situations depicted by the problems, and that each situation falls within an objective category.

Within SECO, the interpretation of the problem statement varies depending on the solver's knowledge: a given situation may thus lead to different encodings. In order to solve an incongruent problem, a solver usually needs to recode the initial representation they have of it. The idea that an initial representation will be recoded to allow the use of a solving algorithm is one that was not covered by Bassok's semantic alignment framework. Bassok and colleagues' theory focuses on the abstraction of an interpreted structure during the initial encoding of a problem (Bassok, 2001), yet what happens when this initial encoding leads to failure hasn't been addressed, especially in cases in which a different representation of the situation could allow the solvers to find the solution. When the first interpreted structure cannot be specified into a valid solving algorithm, SECO covers the possibility that one recodes the situation and manages to solve the problem, in accordance with empirical findings such as the ones reported in Gamo et al. (2010) or Thevenot and Oakhill (2005).

We propose to take a brief look at the empirical prospects opened by SECO. First, because it accounts for the part played by world semantics, SECO predicts that different individuals with different knowledge or experiences about the world may tackle a problem differently. For instance, imagine if Hudson's (1983) problem about birds and worms had been framed in terms of smurfs and mushrooms ("There are 5 smurfs and 3 mushrooms, how many more mushrooms than smurfs are there?"). Children who are familiar with the Smurfs comic series will know that each smurf has his or her own mushroom to live in (there are no housemates in the Smurf village!). Thus, SECO predicts that these children may be more likely to find the solution to the problem, because their world semantics about smurfs and their individual mushrooms will help them to construct a paired encoding in which each house is assigned to one smurf (see Case study 5 for more details on why this should facilitate the solving process).

More generally, SECO makes the prediction that cultural differences in the world semantics evoked by a given problem statement may influence participants' interpreted structure and the subsequent strategies they will use to solve a problem. For instance, it is believed that Indonesian and English speakers tend to represent durations as linear distances (e.g., a *long* time), whereas Spanish and Greek speakers tend to represent durations as definite quantities (e.g., *mucho* tiempo) (Casasanto et al., 2004). Thus, SECO predicts that English and Greek speakers may perform differently on the duration problems used by Gamo et al. (2010) and described in the third case study.

Second, SECO predicts that modifying the world semantics evoked by a problem may influence the interpreted structure encoded. Such representational differences could be measured by asking participants to produce drawings of the problems they solved (e.g., Edens & Potter, 2008). Similarly, recognition tasks may provide a way to probe participants' representation of the problems (e.g., Hegarty, Mayer, & Monk, 1995; Mani & Johnson-Laird, 1982; Verschaffel, 1994), to assess whether their interpreted structures differed depending on the problem statements. Third, a central point in SECO is the recoding pathway, according to which one can recode an interpreted structure into a new representation at a certain cost. This cost can be measured by higher error rates on problems needing a recoding and higher response times on problems successfully recoded (Gros, Sander, & Thibaut, 2019). Future works might even assess the increase in cognitive load associated to this process by measuring physiological responses such as pupil dilation during the recoding of semantically incongruent problems.

Fourth, the existence of the expert encoding pathway may be tested by presenting experts with different problem statements: SECO predicts that experts' performance on problems requiring a recoding may decrease less than that of lay solvers, due to the possibility for experts to directly encode the problems' deep structure, even on incongruent problems. Fifth, SECO accounts for the fact that students may experience difficulty trying to solve a problem if they either lack the relevant world semantics, the relevant mathematical semantics, the ability to recode a semantically incongruent representation or the ability to compute the solving algorithm. Moreover, SECO predicts that different errors will be associated with these different shortcomings. By testing separately students' mathematical knowledge, their world knowledge about the entities described in the problem statement and their ability to compute specific algorithms, SECO can be used to pinpoint and address distinct sources of difficulties.

By providing a finer-grained depiction of solvers' reasoning, SECO can inform future works on the encoding, recoding and solving of arithmetic word problems. The conception of experiments testing the aforementioned predictions should help determine the explanatory power of SECO, either bolstering its claims or leading to the development of new alternative models.

Semantic congruence as an educational lever to tackle arduous notions

The current paper defines semantic congruence and suggests that difficulties might arise when the world semantics evoked by a problem statement is semantically incongruent with the problem's solving algorithm. In this view, semantic incongruence is a source of interference and should be overcome by the learners to efficiently solve the encountered problems. Therefore, developing new methods to help students modulate the influence of world semantics in order to directly access the deep structure of the problems could be especially promising. Still, moderating the influence of world semantics is not trivial, since our knowledge about the world has been shown to be deeply involved in our reasoning, be it relevant or not (Bassok, 2001; Bassok et al., 1998; Gros, Sander, & Thibaut, 2016; Gros, Thibaut, & Sander, 2017).

However, world semantics can also have a facilitative influence. Depending on the semantics attached to a problem, solvers will access a congruent solving algorithm more easily than they would with another problem statement. It has been shown that understanding the situation described in a problem statement can be enough to successfully solve a problem, even for children who did not receive any prior explicit instruction regarding the arithmetic notions required (Carpenter & Moser, 1982; De Corte & Verschaffel, 1987; Ibarra & Lindvall, 1982; Thevenot & Barrouillet, 2015). If the depicted situation is the one "doing the thinking" (Hofstadter & Sander, 2013, p. 432) then the effort is minimal. Depending on the semantics imbued in a situation, its representation might be more or less congruent with the deep structure of the problem and thus render it more or less easy to solve. In this regard, one can imagine that an abstruse mathematical theorem might seem almost self-evident if presented in the appropriate semantic setting.

Designing such situations aiming at fostering the understanding of a complex notion may be achieved through conceptual rewording, as suggested by Vicente et al. (2007). In their study, they highlighted that rewording problem statements in a way that makes more explicit the semantic relations between the problems' entities is beneficial to the solvers. Indeed, difficult problems (i.e., problems that had "to be solved in a different than the actual sequence of the events denoted in the problem," Vicente et al., 2007, p. 837) benefited from conceptual rewording, which referred to situations in which the underlying semantic relations between the given and unknown sets were made more explicit than in the standard version. On the other hand, situational rewording (i.e., when a problem statement is presented in a more enriched and elaborated way, e.g., causal relations between events made more explicit) led to no improvement compared to the standard version.

In SECO's view, conceptual rewording was beneficial because it highlighted the mathematical dependencies between quantities, and thus favored the mapping of the world semantics onto the relevant mathematical semantics. Moreover, simply enriching the semantics of the situation had no effect on the mapping between the statement and mathematical representation. Thus, rewording will work when it aids in building a representation of the mathematical semantics that is congruent with the world semantics.

As a consequence, a crucial application of the SECO model resides in the development of educational interventions treating mathematical learning difficulties by resorting to world semantics in order to help understand and overcome some of the learners' impairments regarding arithmetic understanding. Because SECO differentiates between world semantics, mathematical semantics, and algorithms, it can provide a detailed account of the potential difficulties encountered by students when learning to solve arithmetic word problems. The different components described in the model and the processes that link them are all potential candidates from which specific difficulties may stem. Using SECO, it is possible to differentiate between, for example, a lack of mathematical semantics (e.g., not knowing about the commutative property of multiplication) and difficulties in computing algorithms (e.g., not being able to calculate 3×50), in order to design targeted interventions which would help learners overcome their specific difficulties.

Gaining expertise

One of the distinctive features of the SECO model is that it provides an account of the part played by expertise in the solving of arithmetic word problems. The expert encoding pathway as introduced in SECO accounts for the idea, already developed by Chi et al. (1981), that solvers with sufficient expertise may be able to directly encode the deep structure of a problem, regardless of the world semantics it evokes. Data gathered regarding sorting and solving strategies depending on the learner's level of expertise, in line with Chi et al. (1981) seminal work, provide converging evidence regarding this view (e.g., Schoenfeld & Herrmann, 1982; Silver, 1981). Thus, a crucial educational issue is to promote learners' ability to reach a level of expertise allowing them to directly perceive a problem's deep structure, without first encoding an interpreted structure influenced by their everyday knowledge about the problem's entities.

However, since even expert solvers have been shown to sometimes rely on superficial features to determine their solving strategies (Blessing & Ross, 1996; Novick, 1988), experts' ability to ignore the influence of world semantics in all situations should not be taken for granted. In fact, recent evidence we collected on problems similar to those described in the fourth case study suggests that general expertise in mathematics may not be sufficient to overcome the effects of semantic incongruence (Gros et al., 2019). In this paper, we showed that university-educated adults and expert mathematicians alike were more likely to deem an arithmetic word problem unsolvable when its solution was semantically incongruent with the world semantics evoked by the problem that when the two were semantically congruent.

Does this mean that direct encoding of the deep structure is unattainable? Not necessarily. It could be argued that the influence of world semantics is so pervasive that only specific expertise on the type of problem that is being solved (as compared with general expertise in mathematics) may provide the ability to directly encode the deep structure of the problem. From an educational perspective, the overall goal is to teach students either how to directly perceive the deep structure of the problems they encounter, or at least to efficiently recode an ineffective interpreted structure.

This raises the question of how one may develop such a level of expertise. Although conceptual rewording can be used to make a problem easier to solve, it does not necessarily mean that the solvers will learn to solve other problems which haven't been reworded. Correct answers are worth little if not associated with an increase in expertise. However, deliberately engaging in semantic recoding on multiple occasions on problems sharing the same deep structure may be a path to reach this goal.

In Gamo et al. (2010) study, students' performance improved after they were explicitly told to compare "duration problems" and "number of elements problems," and taught how to semantically recode the number of elements problems to use the 1-step algorithm to solve them. As suggested by the rich literature on deliberate practice (Charness, Tuffiash, Krampe, Reingold, & Vasyukova, 2005; Ericsson, 2004, 2008; Ericsson, Krampe, & Tesch-Römer, 1993; Lehtinen, Hannula-Sormunen, McMullen, & Gruber, 2017; Ward, Hodges, Starkes, & Williams, 2007) repeated training focused on specific tasks such as semantic recoding may be a promising path to develop top-level expertise.

In this perspective, we know ever since Gick and Holyoak's work (1983) on analogical transfer that using different examples describing analogous situations can help represent their common structure (see also Braithwaite & Goldstone, 2015; Kotovsky & Gentner, 1996; Richland, Stigler, & Holyoak, 2012). It thus seems realistic to identify, for any type of problem, which problem statement as well as which sequence of training problems might be the most beneficial to help learners abstract a representation as close to the deep structure as possible. A congruence fading process akin to concreteness fading (Fyfe, McNeil, Son, & Goldstone, 2014) could thus help learners abstract the deep structure of the problems by resorting to increasingly incongruent examples. An interesting venue to capitalize on such effects would be to alternatively present problems attached to different world semantics congruent with different representations, in order to help learners switch from an initial representation to another one, more efficient with regard to the resolution of the problem. Such scaled sequences of problems could be especially efficient if adapted to each learner through the use of Technology Enhanced Learning (Paquette, Léonard, Lundgren-Cayrol, Mihaila, & Gareau, 2006; Shute & Zapata-Rivera, 2012; Tchounikine, 2011). Although these propositions are only hypothetical at this stage, we consider these prospects to be promising leads for conducting further research and for helping foster transfer in mathematics education.

Broader application of the SECO model

An idea at the heart of the SECO model is that the congruence or the incongruence between the world knowledge elicited by a problem statement on one hand and the formal structure of the problem on the other hand can account for solvers' successes and failures, as well as for their need to recode their representations in incongruent situations. We believe this approach can also bear fruits if applied to other educational fields, such as mental arithmetic and non-mathematical problem solving.

Regarding arithmetic non-word problems, studies have shown that embedding an algorithm in a problem statement carrying world semantics may facilitate its computation (Baranes, Perry, & Stigler, 1989; Koedinger, Alibali, & Nathan, 2008; Koedinger & Nathan, 2004; Stern & Lehrndorfer, 1992). SECO details how, depending on the congruence between world semantics and mathematical semantics, the solving process can be either favored or hindered by such an embedment. If an algorithm is embedded in a problem statement carrying congruent world semantics, then finding the solution should be easier.

However, SECO also predicts that a problem statement carrying world semantics incongruent with the algorithm itself should have the opposite effect. Additionally, basic arithmetic operations carry a semantic meaning even when they are not framed within a problem statement (Bell, Swan, & Taylor, 1981; Fischbein, 1989; Fischbein, Deri, Nello, & Marino, 1985; Graeber, Tirosh, & Glover, 1989; Lakoff & Núñez, 2000; Tirosh & Graeber, 1991). According to Fischbein et al. (1985) view, arithmetic operations such as multiplication and division are attached to tacit models imposing constraints on their computation that have no mathematical relevance. For example, they argue that seeing division as the sharing of a collection of objects into a number of equal sub-collections implies that the divisor must be a whole number and that the quotient must be smaller than the dividend.

SECO addresses what happens when the world semantics evoked by the problem statement is incongruent with the objective mathematical structure of the problem. For example, believing that "to divide is to equally share" might lead solvers to rely on semantic knowledge regarding equitable sharing, making it harder to find the solution to arithmetic problems that go against this belief, such as "$8 \div 0.5$." In this view, SECO can guide the analysis of the solvers' activity when faced with such semantic incongruence by showing how the world semantics imbued in the operations themselves evoke an interpreted structure that is incompatible with the solving procedure.

By describing the influence of world semantics on arithmetic problem solving, SECO also underlines the facilitative role that a semantically congruent context may have on arithmetic reasoning in general. Interestingly, the influence of context on the understanding of arithmetic principles has been the focus of several works studying principles such as commutativity or inversion (see Prather & Alibali, 2009, for a review). According to Resnick's (1992, 1994) theory of how mathematical competence is built, arithmetic understanding should emerge following a concrete-to-abstract transition, shifting from an initial object context to a verbal context, then a symbolic context, and then finally to an abstract context. For instance, learning about the commutative property of the addition of two sets of objects may not necessarily mean that learners will immediately be able to transfer this knowledge to the addition of numbers in general (Prather & Alibali, 2009).

In a study about 7- to 9-year-olds' understanding of arithmetic principles, Canobi (2005) showed that some children were helped by a concrete aid to display an understanding of a particular conceptual relation. She showed that some of the participants had an easier time explaining mathematical notions (subtraction complement and inversion principles) when presented with concrete objects instead of abstract numbers. Regarding the principle of commutativity, Cowan and Renton (1996) found that 6- to 9-year-olds showed a better understanding of commutativity in an object context or in a symbolic context, rather than in an abstract context. In other words, performance on mathematically identical tasks depended on the context in which the tasks were presented. As with SECO's description of how the semantic embedding of a word problem can influence learners' ability to find its solution, children's performance in Cowan and Renton's study depended on the context of the task. Finally, in a study on arithmetic problem solving, Jordan, Huttenlocher, and Levine (1992) also found that disparities between middle-income children and low-income children disappeared when the questions were asked using objects rather than when the problems were only posed verbally.

Although few studies have been designed to specifically target the effects of context on principle understanding, and some have reported null effects (e.g., Canobi, Reeve, & Pattison, 2003), most works in the literature are compatible with the theory that children first learn the meaning of arithmetic principles in a grounded context before moving up to higher degrees of abstraction (Prather & Alibali, 2009). A parallel can be drawn with SECO, which accounts for the embedding of an arithmetic problem within a problem statement evoking specific world semantics. In both cases, solvers need to learn how to move away from a grounded encoding and toward a more abstract representation of the situation. We mentioned earlier how the use of increasingly semantically incongruent examples may complement a learning strategy based on concreteness fading (Fyfe et al., 2014), to guide learners from a concrete grasp of a problem to a more abstract understanding of its solution principle. It may be possible to develop a similar strategy in arithmetic learning, by progressively varying the semantic congruence between the concrete situations presented to the learners and the arithmetic notions to be taught.

Regarding problem solving in general, it is well established that the knowledge one has about the entities depicted in a problem can constrain their ability to find a solution (Clement & Richard, 1997; Duncker, 1945; Griggs & Cox, 1982; Kotovsky et al., 1985). Consider, for example, the physics problem consisting in asking whether when a car performs a circular motion at constant speed, its left-side door moves at the same speed as its right-side door or not. Most people trying to solve this problem will use their experience with cars and their world knowledge about rigid objects and represent the two doors of the car as parts of the same object. A common erroneous answer is that when a car moves, every part of the car moves at the same speed since every passenger departs and arrives at the same time.

We believe that the principles underlying SECO can help understand the solvers' reasoning on such a physics problem. In this case, the world semantics used to encode the problem into an interpreted structure will hide some physically relevant aspects of the problem. Unless participants use physics semantics to perform a semantic recoding of the problem that dissociates the two doors as moving along two different circular paths which entails that they do not necessarily travel the same distance, their world semantics will lead them to the erroneous conclusion that the doors travel at the same speed. The notion that congruence between world knowledge and conceptual knowledge associated with a domain of instruction (e.g., mathematical semantics in the case of arithmetic problems, physics semantics for mechanics problems, and so on) can constrain the representation of situations and alter one's reasoning, unless a reinterpretation of the situation happens, seems to be a promising idea. In this regard, the scope of the SECO model could be extended in order to describe the encoding and recoding of situations from different domains of instruction, according to the world semantics and to the domain-related semantics influencing the solvers' interpretation.

Conclusion

The question of how one reasons when solving an arithmetic word problem is a major issue of mathematical education. Understanding the determinants of problem solving is a crucial step in order to identify the difficulties that should be addressed when teaching mathematics. The SECO model provides ground for a distinction between the mathematical semantics, the world semantics, and the algorithms, as well as the way they interact and apply to familiar situations. Those interactions specify the steps involved in the encoding and the recoding of arithmetic word problems.

Being able to foster a semantic recoding in order to improve analogical transfer would be a major step forward in the field of arithmetic teaching, and might help pupils overcome some of their numerous difficulties regarding word problem solving (Gamo et al., 2010; Hegarty, Mayer, & Green, 1992; Richland et al., 2012; Thevenot & Barrouillet, 2015; Verschaffel & De Corte, 1997). Strengthening our grasp of the effects of semantic congruence and incongruence could thus pave the way toward the development of new teaching strategies, building on world and mathematical semantics to guide the students toward a more abstract and more efficient understanding of the encountered problems, contributing to their conception of mathematical notions (Richland et al., 2012).

Acknowledgments

We thank our colleagues from the IDEA and the LEAD labs who provided insight and expertise that greatly assisted the research. We thank Jay Boussier and Katarina Gvozdic for assistance with proofreading.

Funding

This research was supported by the French Ministry of Education and Future Investment Plan under Grant CS-032-15-836-ARITHM-0; French Ministry of Education and Experimental Fund for the Youth under Grant HAP10-CRE-EXPE-S1; Regional Council of Burgundy, Paris Feder Grants under Grants 20159201AAO050S02982 and 20169201AAO050S01845.

ORCID

Hippolyte Gros http://orcid.org/0000-0002-4151-0715
Jean-Pierre Thibaut http://orcid.org/0000-0002-3645-6742
Emmanuel Sander http://orcid.org/0000-0003-0044-3437

References

Baranes, R., Perry, M., & Stigler, J. W. (1989). Activation of real-world knowledge in the solution of word problems. *Cognition and Instruction*, *6*(4), 287–318. doi:10.1207/s1532690xci0604_1

Bassok, M. (2001). Semantic alignments in mathematical word problems. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 401–433). Cambridge, MA: MIT Press.

Bassok, M., Chase, V. M., & Martin, S. A. (1998). Adding apples and oranges: Alignment of semantic and formal knowledge. *Cognitive Psychology*, *35*(2), 99–134. doi:10.1006/cogp.1998.0675

Bassok, M., & Olseth, K. (1995). Object-based representations: Transfer between cases of continuous and discrete models of change. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1522–1538. doi:10.1037/0278-7393.21.6.1522

Bassok, M., Pedigo, S. F., & Oskarsson, A. (2008). Priming addition facts with semantic relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 343–352. doi:10.1037/0278-7393.34.2.343

Bassok, M., Wu, L. L., & Olseth, K. L. (1995). Judging a book by its cover: Interpretative effects of content on problem-solving transfer. *Memory & Cognition*, *23*(3), 354–367. doi:10.3758/BF03197236

Bell, A., Swan, M., & Taylor, G. (1981). Choice of operation in verbal problems with decimal numbers. *Educational Studies in Mathematics*, *12*(4), 399–420. doi:10.1007/BF00308139

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. doi:10.1037/0278-7393.22.3.792

Braithwaite, D. W., & Goldstone, R. L. (2015). Effects of variation and prior knowledge on abstract concept learning. *Cognition and Instruction*, *33*(3), 226–256. doi:10.1080/07370008.2015.1067215

Canobi, K. H. (2005). Children's profiles of addition and subtraction understanding. *Journal of Experimental Child Psychology*, *92*(3), 220–246. doi:10.1016/j.jecp.2005.06.001

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (2003). Patterns of knowledge in children's addition. *Developmental Psychology*, *39*(3), 521–534. doi:10.1037/0012-1649.39.3.521

Carey, S. (2009). *The origin of concepts*. Oxford: Oxford University Press.

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: Lawrence Erlbaum Associates.

Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, I., … Gil, D. (2004). How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish. In K. Forbus, D. Gentner, T. Regier (Eds.), *Proceedings of the 26th annual conference of the Cognitive Science Society* (pp. 575–580). Hillsdale, NJ: Lawrence Erlbaum Associates.

Charness, N., Tuffiash, M., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology*, *19*(2), 151–165. doi:10.1002/acp.1106

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152. doi:10.1207/s15516709cog0502_2

Clement, 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. doi:10.1080/135467897394392

Coquin-Viennot, D., & Moreau, S. (2003). Highlighting the role of the episodic situation model in the solving of arithmetical problems. *European Journal of Psychology of Education*, *3*, 267–279. doi:10.1007/BF03173248

Cowan, R., & Renton, M. (1996). Do they know what they are doing? Children's use of economical addition strategies and knowledge of commutativity. *Educational Psychology*, *16*(4), 407–420. doi:10.1080/0144341960160405

Cummins, D. D. (1991). Children's interpretations of arithmetic word problems. *Cognition and Instruction*, *8*(3), 261–289. doi:10.1207/s1532690xci0803_2

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology*, *20*(4), 405–438. (88)90011-4 doi:10.1016/0010-0285

Davidson, J. E., & Sternberg, R. J. (2003). *The psychology of problem solving*. New York, NY: Cambridge University Press.

Davis-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. *Journal of Educational Psychology*, *83*(1), 61–68. doi:10.1037/0022-0663.83.1.61

De Corte, E., & Verschaffel, L. (1987). The effect of semantic structure on first graders' strategies for solving addition and subtraction word problems. *Journal for Research in Mathematics Education*, *18*(5), 363–381. doi:10.2307/749085

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*(4), 460–470. doi:10.1037/0022-0663.77.4.460

Duncker, K. (1945). On problem-solving (L. S. Lees, Trans). *Psychological Monographs*, *58*(5), i–113. (5, Whole No. 270). doi:10.1037/h0093599

Edens, K., & Potter, E. (2008). How students "unpack" the structure of a word problem: Graphic representations and problem solving. *School Science and Mathematics*, *108*(5), 184–196. doi:10.1111/j.1949-8594.2008.tb17827.x

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*(Supplement), S70–S81. doi:10.1097/00001888-200410001-00022

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine*, *15*(11), 988–994. doi:10.1111/j.1553-2712.2008.00227.x

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. https:// doi:10.1037/0033-295X.100.3.363

Fischbein, E. (1989). Tacit models and mathematical reasoning. *For the Learning of Mathematics*, *9*(2), 9–14.

Fischbein, E., Deri, M., Nello, M. S., & Marino, M. S. (1985). The role of implicit models in solving verbal problems in multiplication and division. *Journal for Research in Mathematics Education*, *16*(1), 3–17. doi:10.2307/748969

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. doi:10.1007/s10648-014-9249-3

Gamo, S., Sander, E., & Richard, J.-F. (2010). Transfer of strategy use by semantic recoding in arithmetic problem solving. *Learning and Instruction*, *20*(5), 400–410. doi:10.1016/j.learninstruc.2009.04.001

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford: Oxford University Press.

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development*, *59*(1), 47–59. doi:10.2307/1130388

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38. doi:10.1016/0010-0285(83)90002-6

Goswami, U., & Brown, A. L. (1990). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition*, *35*(1), 69–95. (90)90037-K doi:10.1016/0010-0277

Graeber, A. O., Tirosh, D., & Glover, R. (1989). Preservice teachers' misconceptions in solving verbal problems in multiplication and division. *Journal for Research in Mathematics Education*, *20*(1), 95–102. doi:10.2307/749100

Greer, B. (1992). Multiplication and division as models of solutions. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 276–295). New York, NY: Macmillan.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, *73*(3), 407–420. doi:10.1111/j.2044-8295.1982.tb01823.x

Gros, H., Sander, E., & Thibaut, J. P. (2016). "This problem has no solution": When closing one of two doors results in failure to access any. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th annual conference of the Cognitive Science Society* (pp. 1271–1276). Austin, TX: Cognitive Science Society.

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. doi:10.3758/s13423-019-01628-3

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 B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 39th annual conference of the Cognitive Science Society*. London, UK: Cognitive Science Society.

Große, C. S., & Renkl, A. (2006). Effects of multiple solution methods in mathematics learning. *Learning and Instruction*, *16*(2), 122–138. doi:10.1016/j.learninstruc.2006.02.001

Guthormsen, A. M., Fisher, K. J., Bassok, M., Osterhout, L., DeWolf, M., & Holyoak, K. J. (2016). Conceptual integration of arithmetic operations with real-world knowledge: Evidence from event-related potentials. *Cognitive Science*, *40*(3), 723–757. doi:10.1111/cogs.12238

Gvozdic, K., & Sander, E. (2017). Solving additive word problems: Intuitive strategies make the difference. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 39th annual conference of the Cognitive Science Society*. London, UK: Cognitive Science Society.

Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology*, *84*(1), 76–84.

Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology*, *87*(1), 18–32. doi:10.1037/0022-0663.87.1.18

Hofstadter, D. R., & Sander, E. (2013). *Surfaces and essences: Analogy as the fuel and fire of thinking*. New York, NY: Basic Books.

Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development*, *54*(1), 84–90. doi:10.2307/1129864

Ibarra, C. G., & Lindvall, C. M. (1982). Factors associated with the ability of kindergarten children to solve simple arithmetic story problems. *The Journal of Educational Research*, *75*(3), 149–156. doi:10.1080/00220671.1982.10885372

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, *4*(1), 71–115. doi:10.1207/s15516709cog0401_4

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (No. 6). Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences*, *107*(43), 18243–18250. doi:10.1073/pnas.1012933107

Jordan, N. C., Huttenlocher, J., & Levine, S. C. (1992). Differential calculation abilities in young children from middle- and low-income families. *Developmental Psychology*, *28*(4), 644–653. doi:10.1037/0012-1649.28.4.644

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, *92*(1), 109–129. doi:10.1037//0033-295X.92.1.109

Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science: A Multidisciplinary Journal*, *32*(2), 366–397. doi:10.1080/03640210701863933

Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. The *Journal of the Learning Sciences*, *13*(2), 129–164. doi:10.1207/s15327809jls1302_1

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*(6), 2797–2822. doi:10.1111/j.1467-8624.1996.tb01889.x

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology*, *17*(2), 248–294. doi:10.1016/0010-0285(85)90009-X

Kouba, V. L. (1989). Children's solution strategies for equivalent set multiplication and division word problems. *Journal for Research in Mathematics Education*, *20*(2), 147–158. doi:10.2307/749279

Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being (Nachdr.)*. New York, NY: Basic Books.

Lehtinen, E., Hannula-Sormunen, M., McMullen, J., & Gruber, H. (2017). Cultivating mathematical skills: From drill-and-practice to deliberate practice. *ZDM*, *49*(4), 625–636. doi:10.1007/s11858-017-0856-6

Leikin, R., & Lev, M. (2007). Multiple solution tasks as a magnifying glass for observation of mathematical creativity. In J. H. Woo, H. C. Lew, K. S. Park, & D. Y. Seo (Eds.), *Proceedings of the 31st international conference for the psychology of mathematics education* (Vol. 3, pp. 161–168). Seoul, Korea: The Korea Society of Educational Studies in Mathematics.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition*, *10*(2), 181–187. doi:10.3758/BF03209220

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. doi:10.1007/BF00366618

Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104, No. 9). Englewood Cliffs, NJ: Prentice-Hall.

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 510. doi:10.1037//0278-7393.14.3.510

Paquette, G., Léonard, M., Lundgren-Cayrol, K., Mihaila, S., & Gareau, D. (2006). Learning design based on graphical knowledge-modeling. *Journal of Educational Technology & Society*, *9*(1), 97–112.

Plato (2009). *Plato's Phaedrus*. Millis, MA: Agora Publications.

Prather, R. W., & Alibali, M. W. (2009). The development of arithmetic principle knowledge: How do we know what learners know? *Developmental Review*, *29*(4), 221–248. doi:10.1016/j.dr.2009.09.001

Resnick, L. B. (1992). From protoquantities to operators: Building mathematical competence on a foundation of everyday knowledge. *Analysis of Arithmetic for Mathematics Teaching*, *19*, 275–323.

Resnick, L. B. (1994). Situated rationalism: Biological and social preparation for learning. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 474–493). Cambridge, UK: Cambridge University Press.

Reusser, K. (1988). Problem solving beyond the logic of things: Contextual effects on understanding and solving word problems. *Instructional Science*, *17*(4), 309–338. doi:10.1007/BF00056219

Reusser, K. (1989). *Textual and situational factors in solving mathematical word problems* (Research Rep. No. 7). Bern, Switzerland: University of Bern, Department of Educational Psychology.

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, NY: Pergamon Press.

Reusser, K. (1993). Tutoring systems and pedagogical theory: Representational tools for understanding, planning, and reflection in problem solving. *Computers as Cognitive Tools*, *1*, 143–177.

Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist*, *47*(3), 189–203. doi:10.1080/00461520.2012.667065

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153–196). New York, NY: Academic Press.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 629. doi:10.1037/0278-7393.13.4.629

Ross, B. H., & Bradshaw, G. L. (1994). Encoding effects of remindings. *Memory & Cognition*, *22*(5), 591–605. doi:10.3758/bf03198398

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.

Sander, E., & Richard, J.-F. (2005). *Analogy and transfer: Encoding the problem at the right level of abstraction*. Proceedings of the 27th Annual Conference of the Cognitive Science Society, Stresa, Italy, pp. 1925–1930.

Schank, R. C. (1975). The role of memory in language processing. In C. N. Cofer (Ed.), *The structure of human memory* (pp. 162–189). San Francisco, CA: Freeman.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans and understanding: An inquiry into human knowledge structures*. Oxford, UK: Lawrence Erlbaum.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(5), 484. doi:10.1037/0278-7393.8.5.484

Shute, V. J., & Zapata-Rivera, D. (2012). Adaptive educational systems. In P. Durlach (Ed.), *Adaptive technologies for training and education* (pp. 7–27). New York, NY: Cambridge University Press.

Silver, E. A. (1981). Recall of mathematical problem information: Solving related problems. *Journal for Research in Mathematics Education*, *12*(1), 54–64. doi:10.3758/10.2307/748658

Squire, S., & Bryant, P. (2002). From sharing to dividing: Young children's understanding of division. *Developmental Science*, *5*(4), 452–466. doi:10.1111/1467-7687.00240

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum.

Staub, F. C., & Reusser, K. (1992). *The role of presentational factors in understanding and solving mathematical word problems*. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

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.

Stern, E., & Lehrndorfer, A. (1992). The role of situational context in solving word problems. *Cognitive Development*, *7*(2), 259–268. doi:10.1016/0885-2014(92)90014-I

Tchounikine, P. (2011). Educational software engineering. In *Computer science and educational software design* (pp. 111–122). Berlin, Heidelberg: Springer.

Thevenot, C. (2010). Arithmetic word problem solving: Evidence for the construction of a mental model. *Acta Psychologica*, *133*(1), 90–95. doi:10.1016/j.actpsy.2009.10.004

Thevenot, C. (2017). Arithmetic word problem solving: The role of prior knowledge. In D. C. Geary, D. B. Berch, R. J. Ochsendorf, & K. M. Koepke (Eds.), *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 47–66). Amsterdam: Elsevier.

Thevenot, C., & Barrouillet, P. (2015). Arithmetic word problem solving and mental representations. In R. Cohen Kadosh & A. Dowker (Eds.), *The Oxford handbook of numerical cognition* (pp. 158–179). New York, NY: Oxford University Press.

Thevenot, C., & Oakhill, J. (2005). The strategic use of alternative representations in arithmetic word problem solving. *The Quarterly Journal of Experimental Psychology Section A*, *58*(7), 1311–1323. doi:10.1080/02724980443000593

Thevenot, C., & Oakhill, J. (2006). Representations and strategies for solving dynamic and static arithmetic word problems: The role of working memory capacities. *European Journal of Cognitive Psychology*, *18*(5), 756–775. doi:10.1080/09541440500412270

Tirosh, D., & Graeber, A. O. (1991). The effect of problem type and common misconceptions on preservice elementary teachers' thinking about division. *School Science and Mathematics*, *91*(4), 157–163. doi:10.1111/j.1949-8594.1991.tb12070.x

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.

Vergnaud, G. (1983). Multiplicative structures. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 127–174). New York, NY: Academic Press.

Verschaffel, L. (1994). Using retelling data to study elementary school children's representations and solutions of compare problems. *Journal for Research in Mathematics Education*, *25*(2), 141–165. doi:10.2307/749506

Verschaffel, L., & De Corte, E. (1997). Teaching realistic mathematical modeling in the elementary school: A teaching experiment with fifth graders. *Journal for Research in Mathematics Education*, *28*(5), 577–601. doi:10.2307/749692

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. doi:10.1348/000709907X178200

Ward, P., Hodges, N. J., Starkes, J. L., & Williams, M. A. (2007). The road to excellence: Deliberate practice and the development of expertise. *High Ability Studies*, *18*(2), 119–153. doi:10.1080/13598130701709715

Supplemental Materials to: *Semantic Congruence in Arithmetic: A New Conceptual Model for Word Problem Solving*

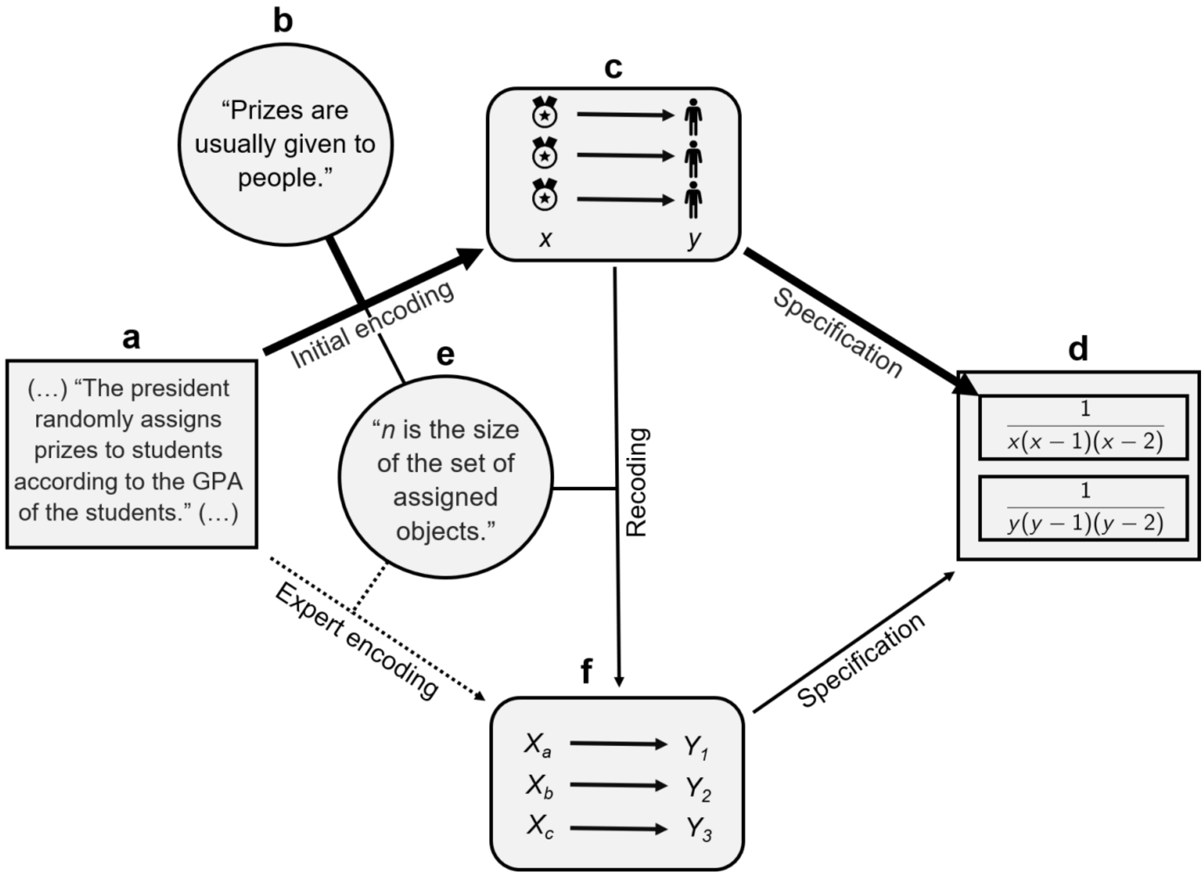

Figure S1. Modeling of the resolution of a permutation transfer problem with an "objects to people" assignment structure, from Bassok, Wu and Olseth (1995).

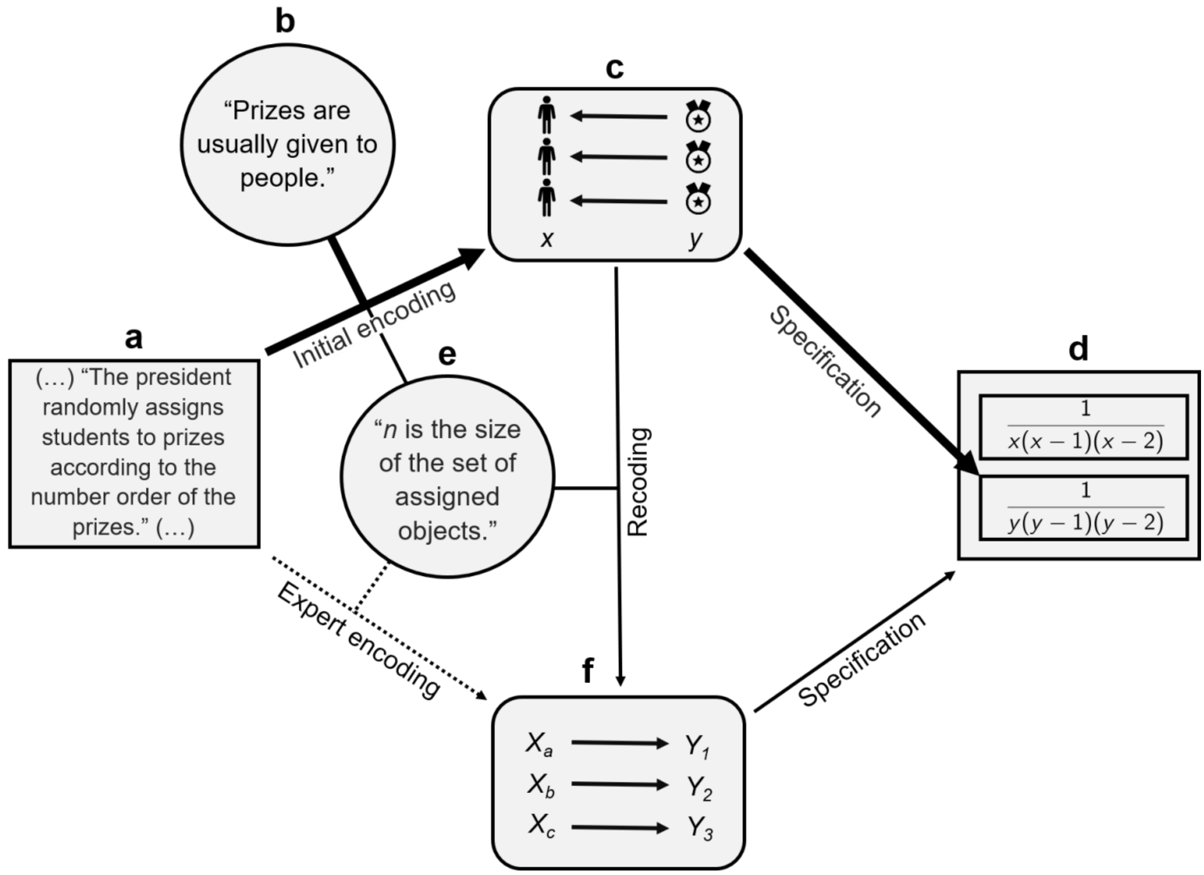

Figure S2. Modeling of the resolution of a permutation transfer problem with a "people to objects" assignment structure, from Bassok, Wu and Olseth (1995).

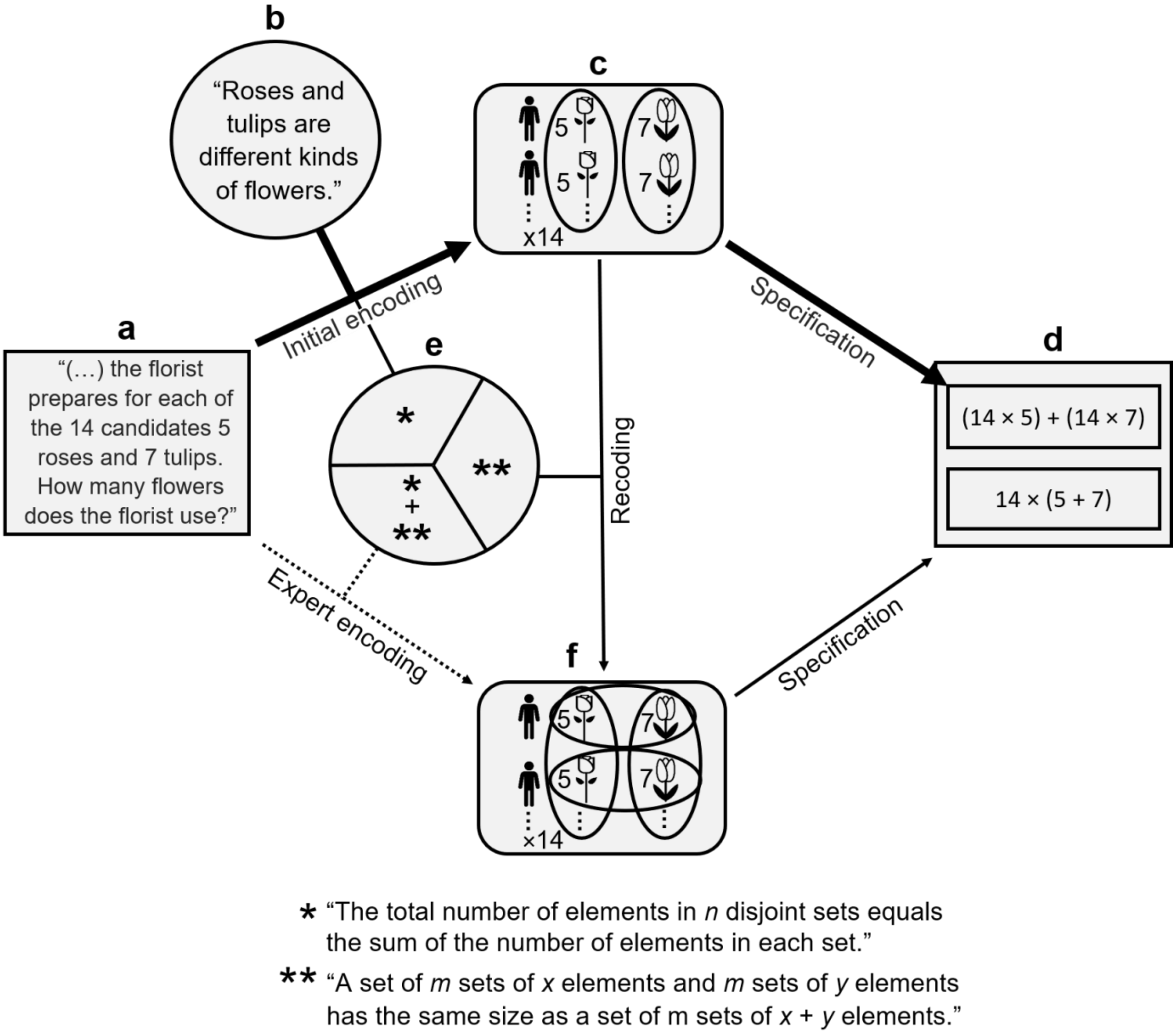

Figure S3. Modeling of the resolution of a distributive problem without a structuring element from Coquin-Viennot and Moreau (2003).

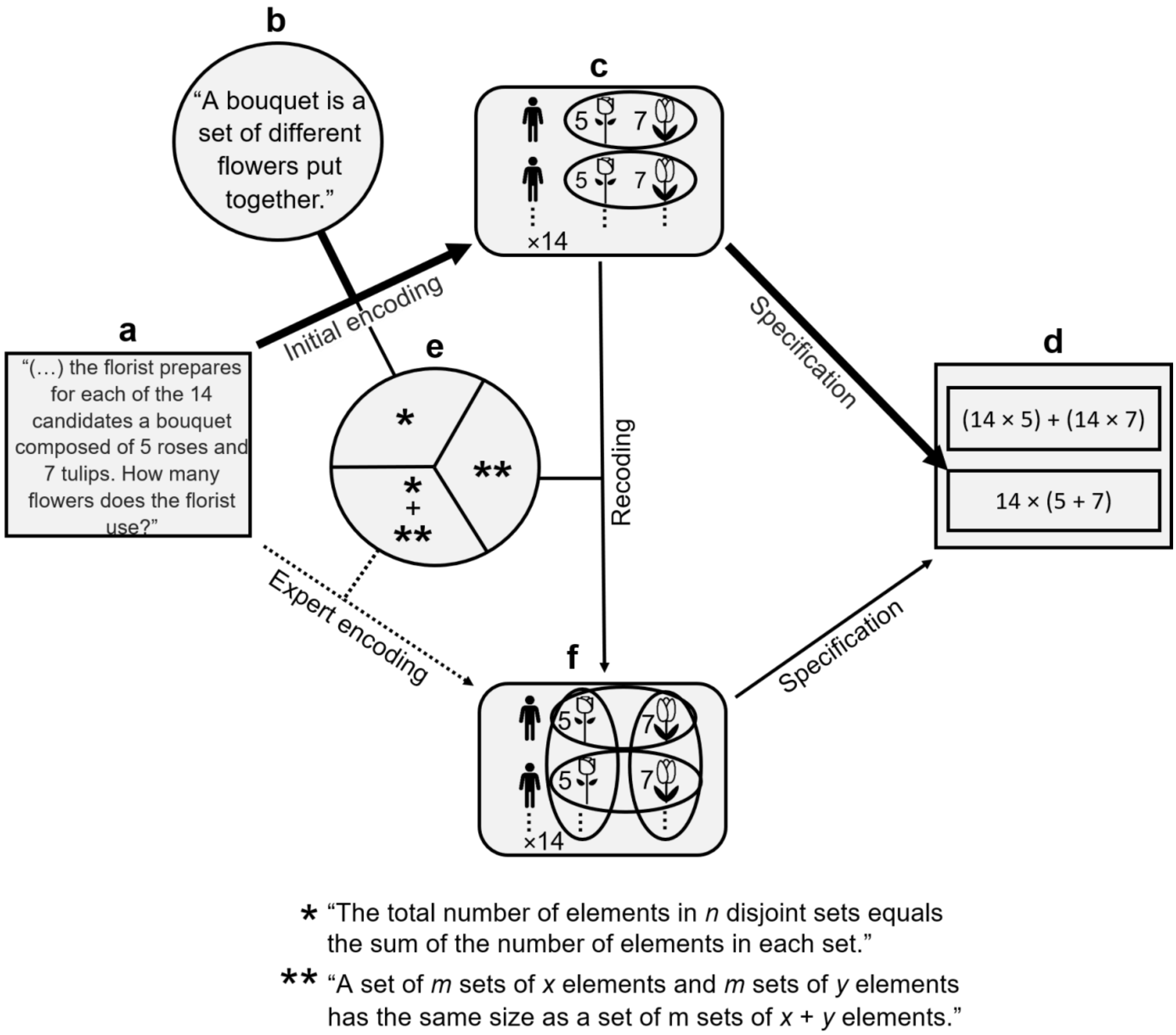

Figure S4. Modeling of the resolution of a distributive problem with a structuring element from Coquin-Viennot and Moreau (2003).

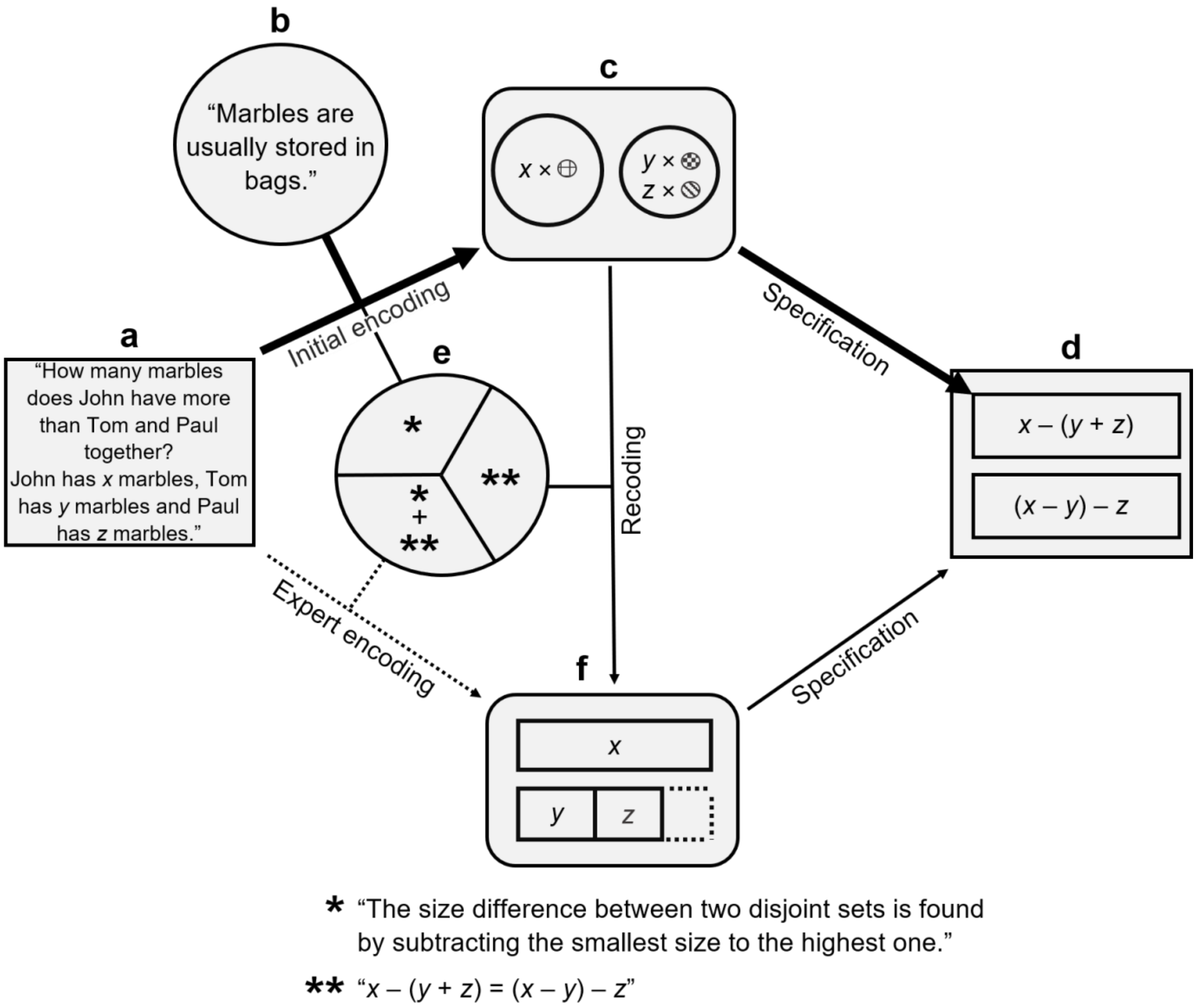

Figure S5. Modeling of the resolution of a "High Cost" problem from Thevenot & Oakhill (2005). This problem could become either a two-digit problem or a three-digit problem depending on the values given to x, y and z.

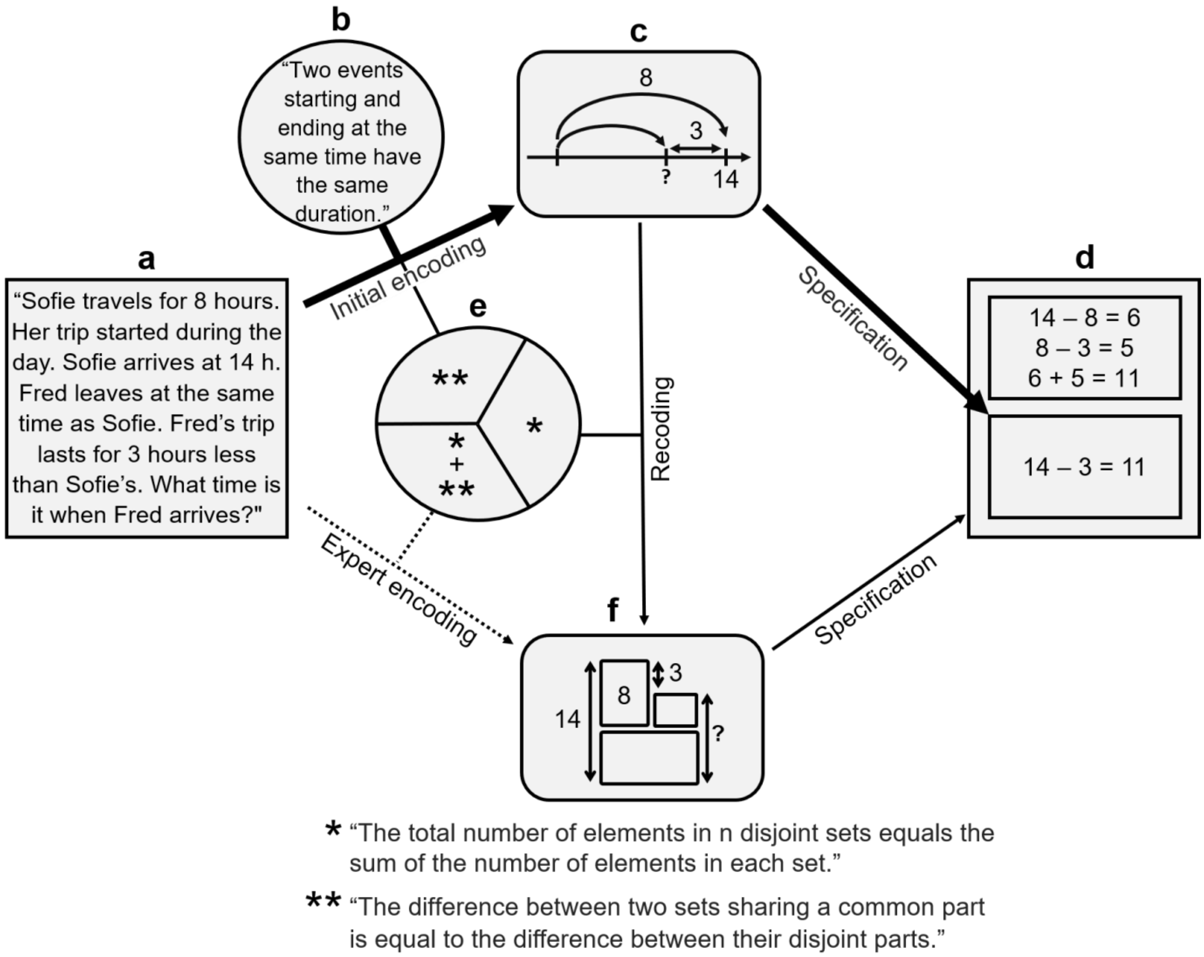

Figure S6. Modeling of the resolution of an ordinal problem from Gamo et al., 2010.

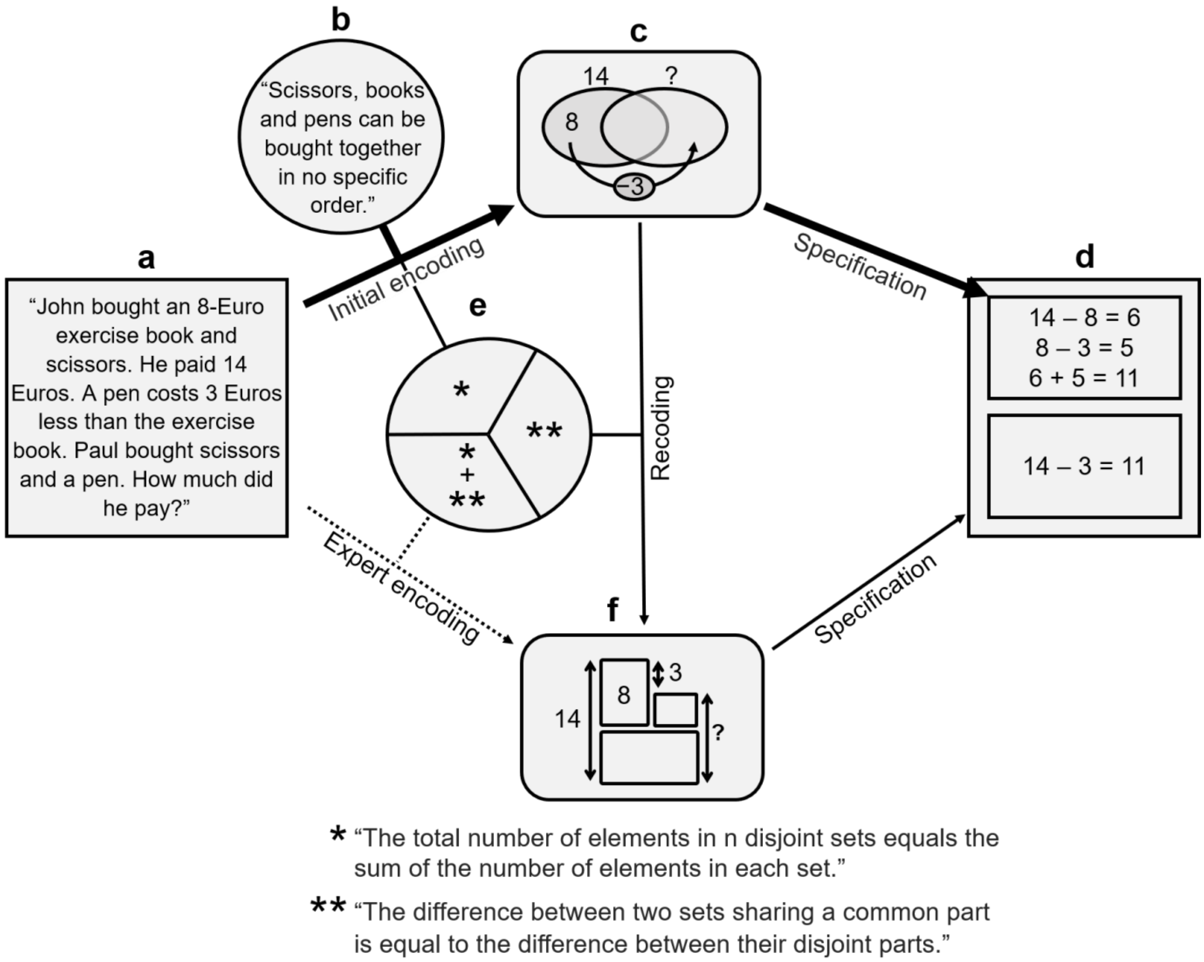

Figure S7. Modeling of the resolution of a cardinal problem from Gamo et al., 2010.

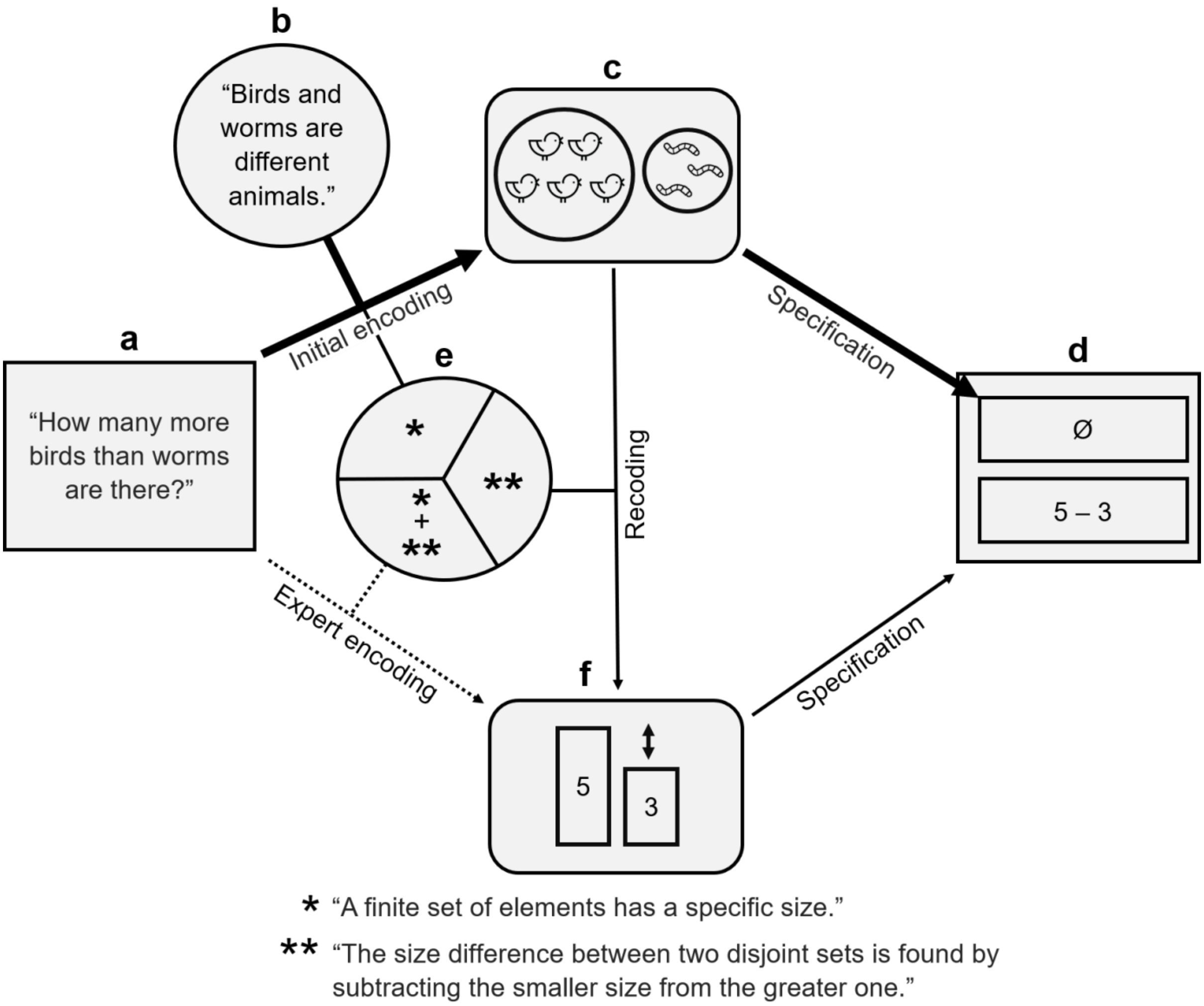

Figure S8. Modeling of the resolution of a "More" problem from Hudson (1983).

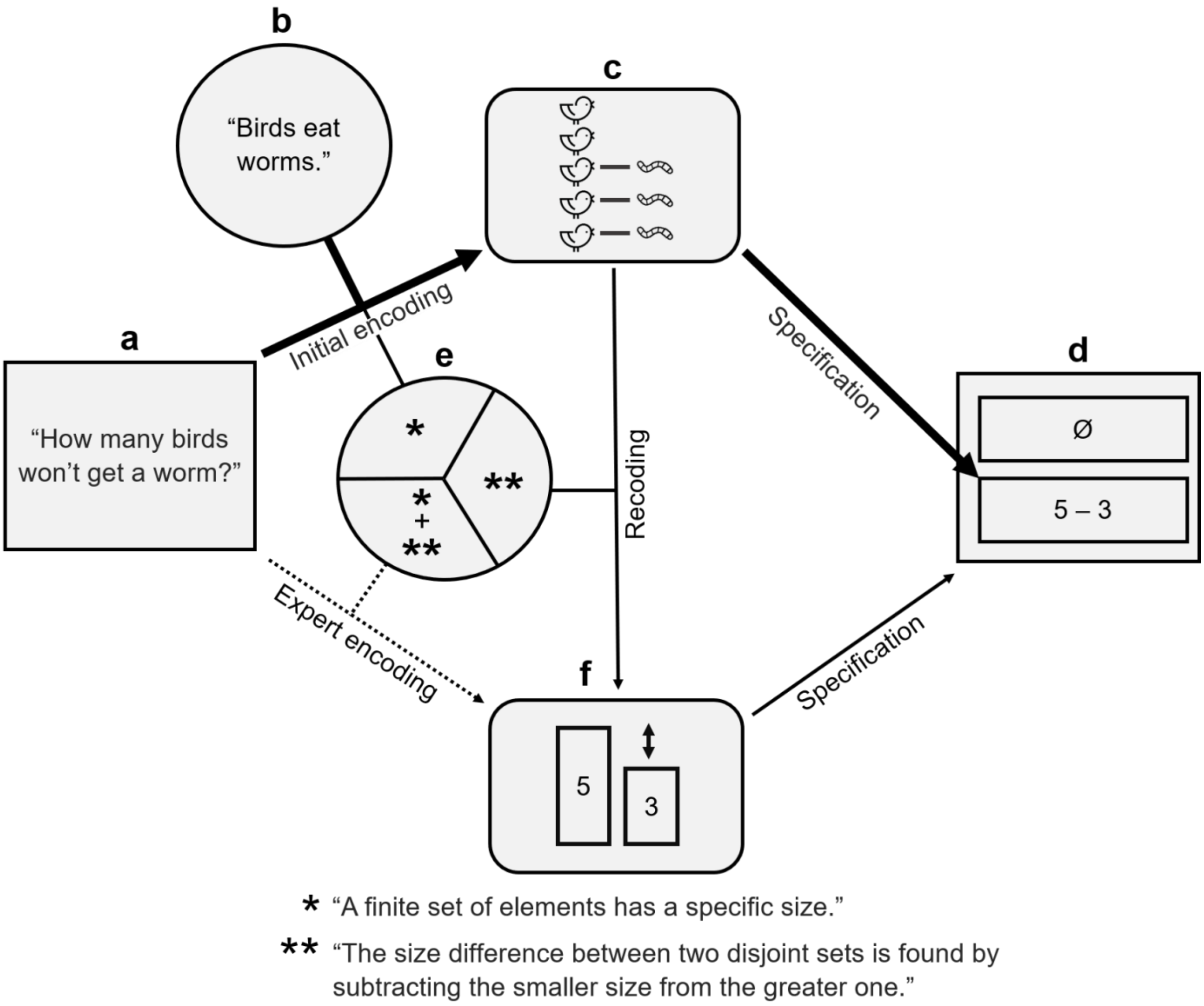

Figure S9. Modeling of the resolution of a "Won't get" problem from Hudson (1983).

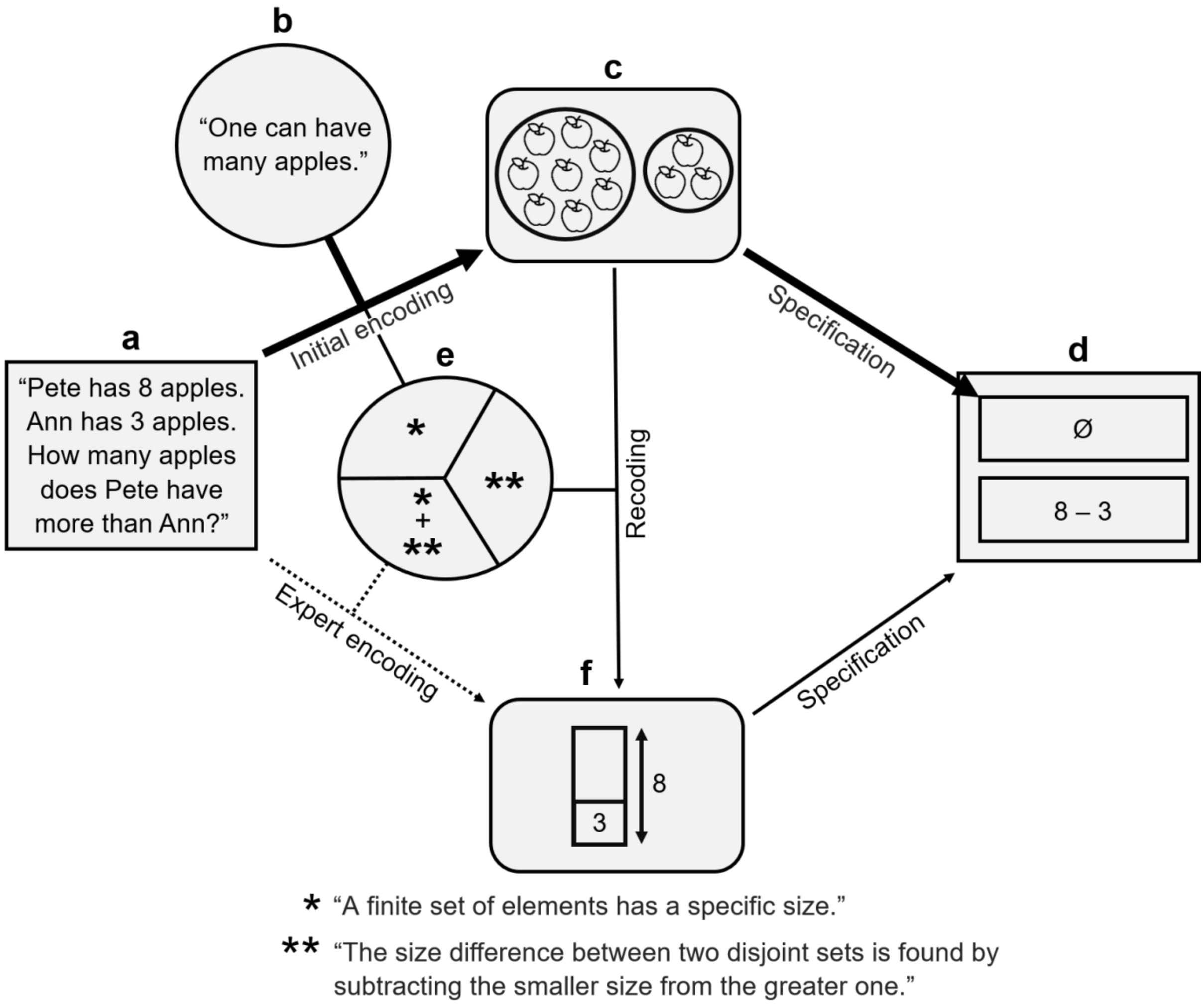

Figure S10. Modeling of the resolution of a standard compare problem from De Corte et al. (1985).

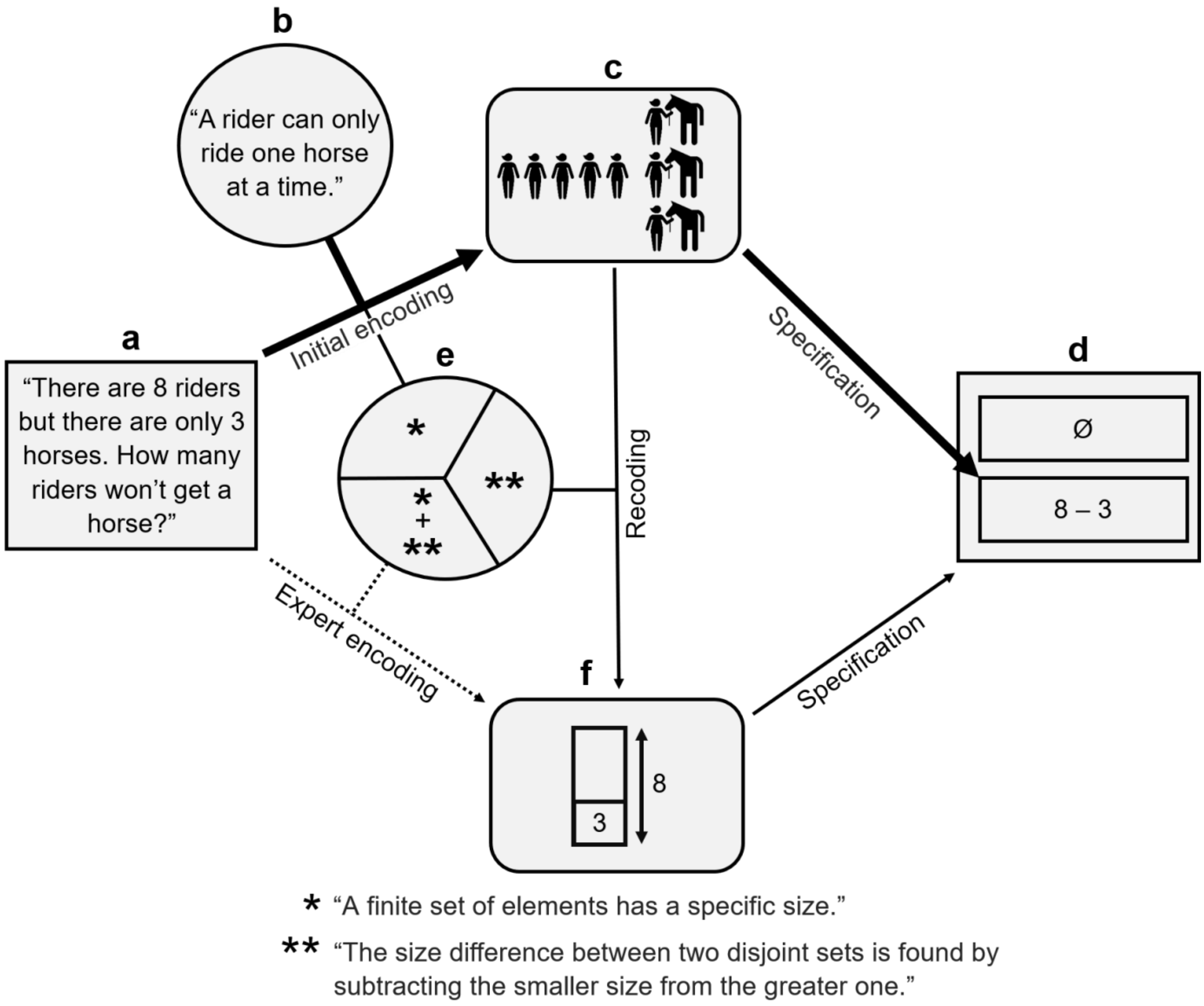

Figure S11. Modeling of the resolution of a reworded compare problem from De Corte et al. (1985).