

Reciprocals and Geometry-sensitivity

An experimental study

Eva B. Poortman

Reciprocals and Geometry-sensitivity: An experimental study

Eva B. Poortman
Student number 3115062
E.B.Poortman@students.uu.nl

MA Linguistics Thesis
Supervisor: Dr. Yoad Winter
Co-supervisor: Dr. Marijn Struiksmā
Third reader: Dr. Joost Zwarts

Utrecht Institute of Linguistics OTS
Utrecht University
June 2011

Abstract

Reciprocal sentences contain a reciprocal expression, a reciprocal antecedent and a relation (e.g. *John, Mary and Sue know each other*). The reciprocal expression *each other* is known to receive a wide variety of interpretations, depending on the predicate in its scope. Previous accounts of reciprocity aim to predict the interpretation of any given reciprocal by selecting a maximal or maximally typical option (e.g. Dalrymple et al., 1998; Sabato, 2006; Kerem, Friedmann & Winter, 2009). In such proposals, maximality is defined in terms of the number of relations among the individuals that make up the antecedent set. This thesis proposes a geometry-sensitive hypothesis, in which the spatial configuration (“geometry”) of those individuals is taken into account as an additional factor.

We provide data for which previous accounts either make incorrect predictions or cannot give a formalized explanation. These data concern reciprocal sentences with two possible interpretations containing unequal amounts of relations and different spatial configurations of individuals (a line and a closed circle). The geometry-sensitive hypothesis that is put forward, determines maximality given a configuration and consequently predicts both of the attested interpretations to be maximal.

Geometry-sensitivity was tested for 22 predicates in an experiment with 71 native speakers of Dutch. We measured acceptability rates of reciprocal sentences in two set-theoretically equivalent situations that differed merely in configuration. The compared situations contained an equal number of individuals (three) and an equal number of relations between them (two), but had the individuals standing either in a line or a circle. Previous accounts that determine maximality only on the basis of number of relations consider the two situations equivalent. For a geometry-sensitive hypothesis, two relations are expected to be maximal given the line configuration but not given the circle configuration.

We found that overall both situations were unacceptable. Geometry-sensitivity was only suggested by results on a subset of the predicates that were tested, namely symmetric predicates. Based on these results, we conclude that geometry-sensitivity does not apply in the broad sense that we initially assumed but possibly in a more restricted way, and we revise our hypothesis accordingly. Finally, we provide suggestions for further research that tests geometry-sensitivity in the revised formulation.

Acknowledgements

Writing this thesis and doing all the preparatory work that it required over the past year has been quite a journey, and it could not have happened without the following people. First of all, big thanks go to my main supervisor Yoad Winter, to whom I am grateful for his enthusiasm and willingness to guide me through the most valuable parts of my master program. All the support and opportunities are very much appreciated. I thank my second supervisor Marijn Struiksma, for kindly helping me out at any time, and in particular for helping me with the statistical analyses of results.

Thanks also to Sophia Katrenko, Assaf Toledo and Hanna de Vries for welcoming me into their research group and commenting on my work during various stages of this thesis. Furthermore, I want to thank Naama Friedmann for going out of her way to discuss my experimental work long distance on many occasions. I give thanks to Nir Kerem, Marja Oudega, Joost Zwarts and the audience at the workshop *On Logic and Common Sense* (June 2011, Utrecht) for helpful suggestions; to all who participated in either one of the pilot studies or the actual experiment; to Ruth Noy Shapira for her work on figures 2, 3, 4, 9, 10, 11, 14 and 16 (plus all drawings in the appendices); and to Gerrit Bloothoof, Sophia Katrenko, Hugo Quené and Yoad Winter for lending me their students for the sake of research. Also, I gratefully acknowledge that the experimental work reported in this thesis was made possible by a VICI grant from the Netherlands Organisation for Scientific Research (NWO).

On a more personal note, I thank my fellow Linguistics students for making the past two years a fun-filled and intellectually rich experience. And last but definitely not least: big hugs to my family and my love, whom I wish to thank for all their love and mental support. I could never have finished this thesis without them. A special acknowledgement goes to the person who makes me feel like there is nothing I cannot do: thanks mom, you're awesome.

Contents

Chapter 1 Introduction	9
Chapter 2 Background on reciprocals	13
2.1 Proposal I: The maximal meaning or interpretation	13
Strongest Meaning Hypothesis (Dalrymple et al., 1998)	13
Maximal Interpretation Hypothesis (Sabato, 2006; Sabato & Winter, 2011).....	17
2.2 Proposal II: The maximally typical interpretation.....	20
Maximal Typicality Hypothesis (Kerem, Friedmann & Winter, 2009).....	20
Chapter 3 Geometry-sensitivity	27
3.1 Circular/linear interpretations and a problem concerning typicality.....	28
3.2 A geometry-sensitive hypothesis	29
3.3 Subconclusion	32
Chapter 4 Experimental study.....	35
4.1 Pilot studies	35
Pilot study 1: Schematically (verbal)	36
Pilot study 2: Picture-based (verbal).....	37
Pilot study 3: Picture-based (written)	39
4.2 The experiment	41
Method	41
Results.....	45
Chapter 5 Discussion.....	49
Chapter 6 Conclusion.....	57
References	59
Appendix A	61
Appendix B.....	73

List of figures

Figure 1 Six logically possible definitions for a reciprocal derived from two parameters (Dalrymple et al., 1998, p.188)	15
Figure 2 Patient cardinality preference: example of visual forced-choice method for <i>to comb</i> (KFW, 2009)	23
Figure 3 Interpretation of reciprocals: example of visual forced-choice method (2 vs. 3 relations) for <i>to comb</i> (KFW, 2009)	23
Figure 4 Interpretation of reciprocals: example of visual forced-choice method (3 vs. 6 relations) for <i>to comb</i> (KFW, 2009)	24
Figure 5 Two situations (schematically)	29
Figure 6 Two configurations (schematically)	30
Figure 7 Possible reference sets according to non-geometry-sensitive vs. geometry-sensitive accounts	32
Figure 8 The two situations that will be compared in the experiment (schematically)	35
Figure 9 Examples of acceptability judgement task (2 items) for pilot study 2	38
Figure 10 Example pair of test items (<i>to clean</i>) in a line and an open circle configuration	43
Figure 11 Example of a question (<i>to comb</i> , line configuration)	44
Figure 12 Outliers (per participant group per order)	47
Figure 13 Outliers (overall)	48
Figure 14 Open circle drawing for <i>to lean on</i>	50
Figure 15 Actual situation (a) and considered alternatives (b)	52
Figure 16 Line and open circle drawings for <i>to touch</i>	53
Figure 17 Line and open circle situations for <i>to be nailed to</i>	54

List of tables

Table 1 Reciprocal meanings attested in natural language (Dalrymple et al., 1998)	14
Table 2 Summary of test items in pilot study 1	36
Table 3 Pilot study 1 results	37
Table 4 Pilot study 2 results	38
Table 5 Summary of test items in pilot study 3	40
Table 6 Pilot study 3 results (total)	40
Table 7 Pilot study 3 results (per verb)	40
Table 8 Lists of verbs tested in the experiment (per questionnaire)	42
Table 9 Acceptability for line and open circle: mean proportion (standard deviation)	45
Table 10 Acceptability for line and open circle per verb: mean proportion	46
Table 11 Compared situations for non-symmetric vs. symmetric verbs	51

Chapter 1

Introduction

Simple reciprocal sentences contain three parts: a reciprocal expression, its antecedent set and a relation. This thesis deals with the semantics of such sentences. An example is in (1).

- (1) John and Mary are pinching each other

In sentence (1), the reciprocal expression *each other* combines the antecedent *John and Mary* and the binary relation *pinch*. Reciprocal sentences that have two members in the reciprocal antecedent set are relatively straightforward to analyse. In the case of sentence (1), we know that in order for the sentence to be true, John must be pinching Mary and Mary must be pinching John. This suggests that a reciprocal expression requires each member of the antecedent set to stand in the given relation to every other member.

Interesting phenomena appear once an antecedent set contains more than two members. Consider examples (2) and (3), with antecedent sets consisting of three members: John, Mary and Sue.

- (2) John, Mary and Sue are standing on each other
(3) John, Mary and Sue know each other

Let us first look at sentence (2). Regardless of the fact that this sentence has the same structure as sentence (1), we cannot generalize the above mentioned rule and assume that each member must stand in the relation *stand on* to each other member to make the sentence true. Sentence (2) obviously does not mean that every person is *standing on* every other person since this is physically impossible. The result is a so called ‘weakening’: the sentence requires merely a situation in which the three individuals stand on each other in a single ‘stack’. By contrast, sentence (3) is an example that does in fact require every person in the antecedent set to stand in the binary relation *know* to every other person. We conclude that sentence (2) has weaker truth conditions than sentence (3), despite the fact that they are structurally identical. Thus, we have seen that the logical interpretation of a reciprocal sentence with three (or more) members in its antecedent set is not as straightforward as one with only two members in its

antecedent set. The interpretation of the reciprocal varies depending on the predicate in its scope.

This interaction between the reciprocal on the one hand and the binary relation on the other hand has received quite some attention among researchers and has led to several models that try to account for the variation of interpretations. Firstly, there are two proposals which are versions of a so-called Strongest Meaning Hypothesis (SMH), in which a semantic principle picks out the maximal meaning for any reciprocal sentence. In Dalrymple et al.'s SMH (1998), the interpretation of a reciprocal sentence corresponds to the strongest meaning - selected from an inventory of possible meanings - that is consistent with known facts about its antecedent, scope and context (p.209). Sabato's (2006) version of the SMH, later referred to as the Maximal Interpretation Hypothesis (MIH) by Sabato and Winter (2011), states that the meaning of a reciprocal sentence is the strongest interpretation that is consistent with the semantic restrictions of the predicate in its scope, without assuming an independent inventory of possible meanings. Secondly, there is a proposal called the Maximal Typicality Hypothesis (MTH) by Kerem, Friedmann and Winter (2009), which does not merely select the strongest possible meaning but looks also at typicality of binary relations as instances of a concept. The meaning of a reciprocal sentence is then proposed to be the one that is maximally typical relative to the predicate concept. In this thesis, we claim that none of the discussed hypotheses, as they are currently formulated, can give a satisfactory account for all data concerning reciprocals. More specifically, we will be dealing with sentences such as (4).

(4) John, Mary and Sue are sitting alongside each other

Sentence (4) is generally judged to be true in a situation in which John, Mary and Sue are sitting alongside each other in a line, while they could also be sitting in a circle, creating a larger amount of "sit alongside-relations". Existing hypotheses either cannot explain why this is the case, or fail to do so in a clear and formalized manner. In this thesis, we aim to explain such facts by proposing to adopt a hypothesis that takes into account a previously ignored parameter, namely geometry. We report on an experimental study that was conducted in order to check this hypothesis.

The structure of this thesis is as follows: Chapter 2 deals with previous works on reciprocals: Dalrymple et al.'s (1998) Strongest Meaning Hypothesis, Sabato and Winter's (2006; 2011) Maximal Interpretation Hypothesis and the Maximal Typicality Hypothesis by Kerem, Friedmann and Winter (2009). Chapter 3 presents potential problematic data for the

accounts so far and introduces a new, geometry-sensitive hypothesis. In chapter 4 we report experimental work aimed at testing this hypothesis. Finally, in chapters 5 and 6 we wrap up the thesis by discussing the implications of the results so far and providing some conclusions.

Chapter 2

Background on reciprocals

2.1 Proposal I: The maximal meaning or interpretation

The first two hypotheses on reciprocals that will be discussed here are Dalrymple et al.'s SMH and Sabato and Winter's MIH, which is based on the SMH. What these two accounts have in common is that they aim to generate a unique set of truth conditions for any reciprocal sentence by looking only at definitional aspects of the meaning of predicate concepts.

Strongest Meaning Hypothesis (Dalrymple et al., 1998)

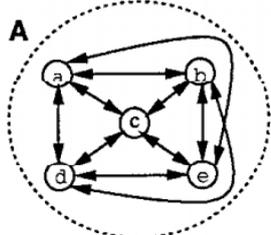
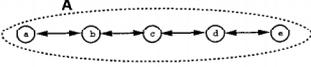
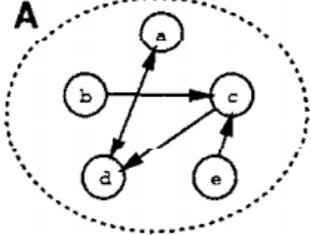
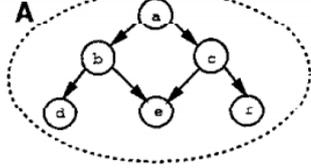
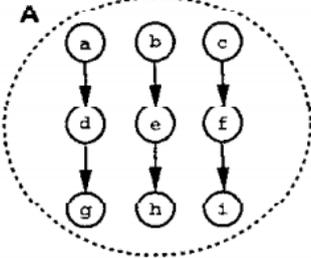
In order to account for the variation of meanings of reciprocal sentences, Dalrymple et al. (1998) propose two separate components: 1) an inventory of all the possible meanings for reciprocal sentences and 2) the SMH, a principle that selects one candidate from these possibilities. Dalrymple et al. start off by exploring the meaning of *each other* and observe that “the interpretation of *each other* varies in meaning according to the meaning of their scope and antecedent as well as the context in which they are uttered” (p.159). In their paper, they make a taxonomy of reciprocal meanings that are claimed to be attested in natural language¹ and then derive parameters of variation that give us six different logically possible definitions of reciprocity. Their final step is taking the three components that appear to be relevant in determining a reciprocal's interpretation (scope, antecedent and non-linguistic information), and combining them into a principle that picks out one meaning for any reciprocal sentence.

Taxonomy and parameterization Dalrymple et al. (1998) distinguish five reciprocal meanings in natural language and give examples for each one. These are summarized in table 1 below. The meanings are ordered with respect to strength, such that the one at the top of the table (Strong Reciprocity) is the logically strongest among the five (non-equivalent) meanings.

¹ It is unclear what Dalrymple et al.'s criterion for attesting a meaning is. Here we merely refer to the reciprocal meanings that are claimed to be attested by Dalrymple et al. in their paper.

As in Dalrymple et al.'s convention, we use A for the group denoted by the antecedent of the reciprocal and R for the relation that holds between them.

Table 1 Reciprocal meanings attested in natural language² (Dalrymple et al., 1998)

<i>Reciprocal meaning</i>	<i>Formal definition</i>	<i>Informal definition</i>	<i>Example</i>
Strong Reciprocity ³	$\forall x, y \in A (x \neq y \rightarrow Rxy)$	Every member of A is related directly by R to every other member	
Intermediate Reciprocity ⁴	$\forall x, y \in A (x \neq y \rightarrow \text{for some sequence } z_0, \dots, z_m \in A (x = z_0 \wedge Rz_0z_1 \wedge \dots \wedge Rz_{m-1}z_m \wedge z_m = y))$	Every member of A is related directly or indirectly by R to every other member	
One-way Weak Reciprocity	$\forall x \in A \exists y \in A (x \neq y \wedge Rxy)$	Every member of A participates with some other member in the relation R as the first argument of the relation	
Intermediate Alternative Reciprocity	$\forall x, y \in A (x \neq y \rightarrow \text{for some sequence } z_0, \dots, z_m \in A (x = z_0 \wedge (Rz_0z_1 \vee Rz_1z_0) \wedge \dots \wedge (Rz_{m-1}z_m \vee Rz_mz_{m-1}) \wedge z_m = y))$	Every member of A is connected directly or indirectly by R, ignoring the direction	
Inclusive Alternative Ordering ⁵	$\forall x \in A \exists y \in A (x \neq y \wedge (Rxy \vee Ryx))$	Every member of A participates with some other member in the relation R as the first or as the second argument (not necessarily both)	

² Again, whether a reciprocal meaning is attested is a theoretical claim made by Dalrymple et al. (see also footnote 1)

³ Strong Reciprocity was originally formulated by Langendoen (1978)

⁴ Intermediate Reciprocity was originally formulated by Langendoen (1978)

⁵ Inclusive Alternative Ordering was originally formulated by Kański (1987)

Now that we have given Dalrymple et al.'s taxonomy of reciprocal meanings in natural language, we move on to the parameters of variation. Reciprocals are $\langle 1,2 \rangle$ quantifiers that take a set and a binary relation and return a truth value (Dalrymple et al., 1998, p.183). Two parameters formulated by Dalrymple et al. give rise to six possible truth conditions for the reciprocal. Five of these are attested in natural language (see table 1), for the sixth possibility they do not provide an example but claim that it might also turn out to be an actual meaning.

The first parameter of variation concerns how the scope relation R should cover the domain A. For this parameter there are three options: 1) each pair of individuals in A participates in the relation R directly (FUL), 2) each pair of individuals in A participates in the relation R either directly or indirectly (LIN), and 3) each single individual in A participates in the relation R with another one (TOT). Moreover, the authors make the more general remarks that pairs with two identical members (identity relations I) have no influence on the meaning of a reciprocal sentence, and that the set A must consist of at least two members (Dalrymple et al., 1998, p.186-187).

The second parameter of variation has to do with how the reciprocal's scope determines the argument R. For this parameter there are two possibilities: 1) the argument is the relation denoted by the predicate R, and 2) the argument is obtained from the relation by ignoring the direction of that relation. So the reciprocal applies either to R or to R's symmetric closure (R^\vee). When we combine parameter 1 and parameter 2, with three and two possibilities respectively, we arrive at 6 (3×2) logically possible meanings for a reciprocal. The overview is given in figure 1 below, taken from Dalrymple et al. (1998). In this figure, we can see that Strong Alternative Reciprocity (SAR) is the sixth, so far unattested, logical possibility. SAR has the parameter settings FUL and R^\vee , thus its meaning is such that each pair of individuals in A participates in the relation R directly and the reciprocal applies to R's symmetric closure ($\forall x,y \in A (x \neq y \rightarrow (Rxy \vee Ryx))$).

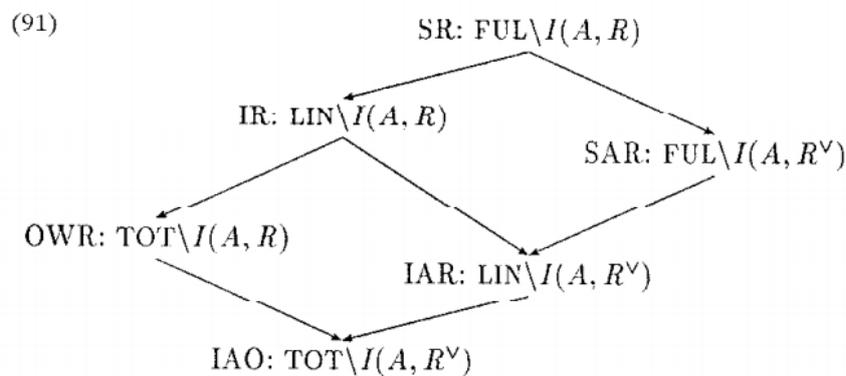


Figure 1 Six logically possible definitions for a reciprocal derived from two parameters (Dalrymple et al., 1998, p.188)

Selecting one candidate As already mentioned, Dalrymple et al. (1998) claim that the meaning of *each other* is flexible or context-sensitive in the sense that it varies depending on its scope, antecedent and further linguistic and non-linguistic information. This means that the meaning of the reciprocal needs to be consistent with properties of the relation (expressed by the parameters above), the domain it ranges over and the context in which it is uttered. Dalrymple et al. combine these three factors into the SMH in order to arrive at a single context-sensitive meaning for each reciprocal:

Strongest Meaning Hypothesis: A reciprocal sentence S can be used felicitously in a context c , which supplies non-linguistic information I relevant to the reciprocal's interpretation, provided the set X_c has a member that entails every other one:

$$X_c = \{ p \mid p \text{ is consistent with } I \text{ and } p \text{ is an interpretation of } S \text{ obtained by interpreting the reciprocal as one of the six quantifiers in figure 1} \}$$

In that case, the use of S in c expresses the logically strongest proposition in X_c (Dalrymple et al., 1998, p.193).

According to this principle, the parameters of reciprocal meaning are set so that the reciprocal sentence as a whole has the strongest possible candidate meaning. Let us make this more concrete by looking at a few examples. Consider sentence (5), an example from Dalrymple et al. (1998).

(5) Five Boston pitchers sat alongside each other

The interpretation of sentence (5) is obviously not such that every pitcher is sitting next to every other pitcher (since this is physically impossible), but it is such that the pitchers are sitting alongside each other in a line. This means that the reciprocal is interpreted as Intermediate Reciprocity (IR): every member of the set of pitchers is related directly or indirectly by the relation *sit alongside* to every other member. Let us look at the steps that Dalrymple et al. assume in order to get to this interpretation. First considering the scope of the reciprocal, Dalrymple et al. note that the relation *sit alongside* is symmetric. This characteristic alone does not exclude any of the proposed meanings. However, when we also consider the facts that 1) the antecedent set contains five individuals and 2) every individual has only two sides and can thus only sit next to two other individuals, SR and SAR are ruled out. We are left with four possible meanings that are consistent with the requirements of the

relation, antecedent set and context: IR, IAR, IAO and OWR. Out of these remaining possibilities, the SMH selects IR since it is the strongest among them. Another example is in (6).

- (6) They stacked tables on top of each other (in order to climb through a window)

According to Dalrymple et al. (1998; p.191), the interpretation of the reciprocal in sentence (6) is Intermediate Alternative Reciprocity: there is one stack of tables. In contrast to the relation in (5), here we are obviously dealing with an asymmetric relation, which already excludes SR. Next, considering the fact that the antecedent set (the number of tables) must be finite, one table must be at the bottom and thus not participating in the relation as the first argument: it is not stacked on top of another table. This fact excludes both OWR and IR. Based on the same information we can also exclude SAR, since it requires that every pair is in direct contact, which is impossible in case there are more than two tables. We are now left with IAR and IAO, in this case meaning that there is one stack of tables (IAR) or that multiple stacks are possible (IAO). The SMH selects the strongest possibility that is left and that is consistent with the context, and this is IAR. To conclude, the reciprocal is always interpreted so as to “maximize the strength of the combined meaning of the reciprocal’s scope and antecedent” (p.193).

Maximal Interpretation Hypothesis (Sabato, 2006; Sabato & Winter, 2011)

Another system that aims to predict the interpretation of reciprocal sentences is Sabato’s version of the SMH, which will be referred to as the Maximal Interpretation Hypothesis (MIH) as is done in Sabato & Winter (2011). The main difference compared to Dalrymple et al.’s SMH is that the MIH does not consist of two separate components. Whereas Dalrymple et al.’s proposal is composed of the SMH on the one hand and a set of possible meanings that the SMH chooses from on the other hand, Sabato’s system does not make use of an independent set of meanings. Instead, the interpretation of the reciprocal is derived immediately from the properties of the predicate in its scope (Sabato, 2006; Sabato & Winter, 2011).

Semantic restrictions on predicates The MIH functions as “a mapping from semantic restrictions on the predicate’s denotation into the interpretation of the reciprocal, without independent assumptions about available reciprocal meanings” (Sabato, 2006, p.3). This means

that 1) the interpretation of a reciprocal expression is not chosen from an independent set of possible meanings for all reciprocals, and 2) the interpretation is not influenced by ‘contextual information’ in the way that Dalrymple et al. (1998) describe. Instead, it is influenced by some restrictions inherent to a predicate which limit the range of a priori possible interpretations (with no assumptions about which of these are actually attested in natural language (Sabato, 2006, p.25)). Sabato (2006) gives arguments against Dalrymple et al.’s notion of context: she observes that not all contextual information leads to weakening of the reciprocal meaning, thus this notion is too broad to explain the facts. Consider the sentences in (7), an example taken from Sabato (2006).

- (7) #John and Bill don’t know each other. John, Bill and Dan know each other.

According to Dalrymple et al.’s SMH, the sentences in (7) should not be contradictory since the first sentence would function as a context for the second sentence, thereby causing the reciprocal to receive a weaker interpretation. In fact, this is not what happens and we consider (7) to be very odd. Instead of using a vague notion of context, Sabato restricts ‘context’ to semantic properties of the predicate, called ‘semantic restrictions’. This means that in determining the interpretation of a reciprocal, her system only considers the collection of relations (or ‘graphs’) that are possible as denotations of the predicate in the first place. Sabato (2006) calls this subset of all possible graphs ‘the reciprocal interpretation domain’. Let’s make the notion of semantic restrictions more clear by looking at some examples that Sabato provides, reproduced here as (8) and (9).

- (8) These three people like each other
(9) The 3rd grade students gave each other measles

The predicate *like* in sentence (8) has no semantic restrictions. This means that it may denote any binary relation and thus its reciprocal interpretation domain contains all possible graphs. In (9) on the other hand, the predicate *give measles* is restricted in such a way that it can only be the inverse of a function (a person cannot get measles twice) and can only denote acyclic relations (measles cannot be passed around in a circle) (Sabato, 2006, p.33). Consequently, the graphs that are considered as possible interpretations for this reciprocal are only the ones that are consistent with these restrictions.

The MIH as a local maximality principle

Now that we have discussed the notion of semantic restrictions, we move on to the role of the MIH according to Sabato. She claims that the MIH needs to be implemented as a so-called local maximality principle.

Maximal Interpretation Hypothesis: A reciprocal sentence is consistent with models in which no pairs of non-identical individuals in the antecedent set can be added to the denotation of the predicate within its semantic restrictions (Sabato, 2006, p. 31).

The principle refers to maximality because it picks out a single, maximal interpretation for any reciprocal sentence. It is local with respect to the way that the evaluation works: “to compute whether a relation is a local maximum, it is enough to consider configurations that result from adding pairs to the examined relation” (p.32). This kind of evaluation stands in contrast to global maximality, in which it is necessary to consider all other possible configurations in order to compute whether a relation is maximal. Sabato (2006) claims that a local kind of evaluation mechanism is preferred for both cognitive and empirical reasons (p.31-32). Unlike Dalrymple et al’s SMH which considers all possibilities when selecting the strongest candidate, the MIH looks at a specific situation and checks merely whether pairs can be added to that specific case or not (while staying within the reciprocal interpretation domain) – thereby sort of reaching the maximal ‘one step at a time’.

We can now combine our knowledge of semantic restrictions and this MIH as a local maximality principle in order to predict interpretations of actual reciprocal sentences. Let’s again consider sentence (8) and (9). As already mentioned, the relation *like* in (8) contains no restrictions so that it may denote any binary relation – creating the largest possible reciprocal interpretation domain. Now the MIH says that the correct interpretation of this sentence is the one in which no pairs of non-identical individuals can be added, while staying consistent with the semantic restrictions. Since in this case there are no restrictions, the interpretation of the sentence is such that each person in the antecedent set of three people likes each other person (SR). This is the only interpretation where no other pairs can be added. As we saw, the relation *give measles* in (9) does have semantic restrictions, thus it has a smaller interpretation domain. The MIH selects the strongest one among these, resulting in an interpretation where “every 3rd grade student is connected to every other student by the transitive and symmetric closure of the denotation of *give measles*” (p.34). Such an interpretation is maximal in the sense that no more pairs of non-identical members can be added, given the restrictions

inherent to the predicate concept. Thus, we have seen that from properties of the predicate itself, we can reach truth conditions of the reciprocal sentence as a whole. Note that this way Sabato simply does not need an independent set of possible meanings for a reciprocal, since its interpretation is determined by each predicate.

To conclude, the MIH is based on Dalrymple et al's SMH in the sense that it is a principle aimed at predicting the interpretation of a reciprocal sentence by evaluating and choosing the strongest candidate. The difference is that this new system derives the interpretation directly, by having the MIH operate on semantic restrictions of the predicate in the reciprocal's scope. Thus it adds the interpretation domain of the predicate as an extra argument of the reciprocal. Consequently, unlike in Dalrymple et al.'s SMH where each reciprocal expression has exactly the same interpretation domain (see figure 1), in the MIH the interpretation domain of the predicate is the domain of the reciprocal function (Sabato, 2006; Sabato & Winter, 2011). This domain for each reciprocal is a (proper) subset of the total domain and this gives Sabato's system the advantage of being cognitively simpler.

2.2 Proposal II: The maximally typical interpretation

The two accounts that were described in section 2.1 are systems that generate truth conditions for a reciprocal sentence by selecting the maximal candidate. They assume a definitional, Classical Theory of concepts, in which an entity either is or is not an instance of a concept. Another type of proposal takes such a two-valued treatment of concepts to be insufficient and deals not only with truth and falsity but takes into account typicality preferences. Kerem, Friedmann and Winter (2009) developed such a proposal using a principle called the Maximal Typicality Hypothesis (MTH). According to this hypothesis, the interpretation of a reciprocal expression is determined not using its maximal meaning or interpretation, but rather using its most typical one.

Maximal Typicality Hypothesis (Kerem, Friedmann & Winter, 2009)

Kerem, Friedmann and Winter (2009) (henceforth KFW) address the same problem as the two accounts described above, namely the fact that reciprocal expressions do not always make the same logical contribution to sentences, leading to a variety of possible interpretations. The authors argue against previous accounts, underlining the need to look at typicality preferences

in order to understand the many interpretations of reciprocals. As already mentioned, both the SMH and the MIH are only sensitive to so-called definitional aspects of meaning. KFW (2009) point out a specific case where this sort of approach makes incorrect predictions, given in (10).

- (10) Larry, Monty and Garfield are combing each other

The SMH predicts that there is one interpretation of this sentence, namely such that each man is combing every other man (6 “comb-relations” in total), and that all other possible situations make this sentence false. Dalrymple et al.’s argumentation would be that this is the maximal interpretation since it is the strongest candidate that is consistent with antecedent, scope and context. The MIH would also predict that this is the maximal one, since this interpretation is within the domain of the predicate and no pairs of non-identical individuals can be added. What KFW (2009) claim, and give experimental evidence for, is that this interpretation does not block weaker interpretations. In fact, a situation with 3 “comb-relations” in total, where each man combs only one other man, is even preferred over the strongest meaning. Such an example is a clear counter argument for both the original SMH and the derived MIH.

Concepts and typicality What KFW’s account is based on is a theory of concepts other than the Classical Theory. Theories of concepts deal with how concepts are represented in the mind and how we categorize objects as instances of certain concepts. In the Classical Theory, concepts have definitional structure such that every concept encodes necessary and sufficient conditions (Laurence & Margolis, 1999, p.10). For example, the concept ‘bachelor’ has the properties ‘male’, ‘adult’ and ‘unmarried’. Each of these properties are necessary conditions in order to be counted as a bachelor, and taken together they are (arguably) sufficient to define the concept ‘bachelor’. Categorization is straightforward in the Classical Theory: an entity either does or does not fall into the category that instantiates a concept. However, a whole range of psychological studies have shown that such a theory fails to hold in reality for one-place predicates (e.g. Rosch, 1973; Smith, Shoben & Rips, 1974; Rosch & Mervis, 1975). Concepts give rise to typicality effects: membership within a category appears to be graded. For example, subjects consistently judge a ‘robin’ as a more typical bird than a ‘penguin’, even though both contain sufficient properties for the concept ‘bird’ (e.g. Rosch, 1973). Such typicality effects have been shown using many different dependent variables, such as typicality ranking, categorization speed and error rates.

A theory that does take into account typicality is Prototype Theory. According to Prototype Theory, concepts are “structured mental representations that encode the properties that objects in their extension *tend* to possess” (Laurence & Margolis, 1999, p.31). This means that an entity need not satisfy all necessary conditions in order to be counted as an instance of a certain concept, but a sufficient number of them. In Prototype Theory, the categorization process works via similarity comparison between an entity and a concept’s prototype (a collection of perceptually salient properties). This way, an entity can be more or less similar to the prototype, leading to graded exemplariness. Kerem, Friedmann and Winter (2009) claim that any theory that deals with concepts should take typicality preferences into account. Thus they believe that this type of gradedness is not only applicable to one-place predicates, but plays a crucial role with binary predicate concepts as well. More precisely, what the authors claim is that typicality of binary relations is “the core semantic information on binary predicate concepts” (KFW, 2009, p.7) and thus plays a role when a predicate combines with a reciprocal expression. They propose a new hypothesis, which acknowledges typicality:

Maximal Typicality Hypothesis: A reciprocal expression requires the denotation of its predicate antecedent to be a relation of maximal typicality relative to the predicate concept (Kerem, Friedmann & Winter, 2009, p.3).

According to the above hypothesis, a reciprocal expression need not receive the maximally possible interpretation but instead receives the maximally typical one. Looking back at sentence (10), KFW would explain the facts by stating that it is more *typical* for the concept ‘comb’ to comb one other person at a time instead of two. This is exactly what they show in a series of experiments, which will be discussed next.

Experimental evidence The MTH predicts a relation between typicality effects and the interpretation of reciprocal expressions. Any study that aims to test the MTH thus requires pairs of experiments, where one tests typicality for predicates and the other tests interpretations of reciprocal sentences containing those predicates. KFW tested the predictions made by the MTH using two such pairs: two visual experiments and two textual experiments.

In order to test typicality preferences for binary predicates, KFW used the parameter ‘patient cardinality’. The visual experiment was a forced-choice picture selection task. Participants were presented with items consisting of two illustrations, namely an agent acting

on one patient vs. an agent acting on two patients (see figure 2), and a sentence without an object (e.g. *The boy is combing*). They were then asked which of the two depicted situations better describes the sentence. In the textual experiment, participants were given incomplete sentences and were asked to complete the sentence by adding either a singular or plural object (see example (11)).

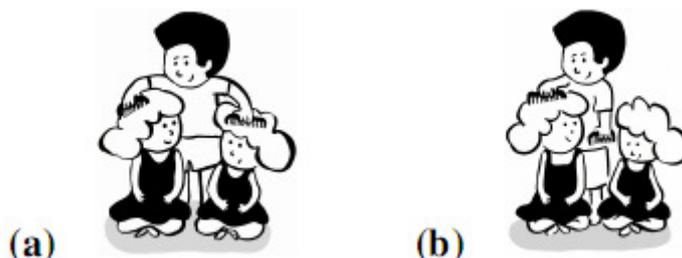


Figure 2 Patient cardinality preference: example of visual forced-choice method for *to comb* (KFW, 2009)

(11) Example of textual elicited judgement task for *to compliment* (KFW, 2009)

ba-Sana	Se'avra	Omer hexmi	a. la-yalda	b. la-yeladot
in-the-year	that-passed	Omer complimented	a. to-the-girl	b. to-the-girls

“Last year, Omer complimented... a. the girl b. the girls”

For testing interpretations of reciprocal expressions, KFW again conducted a forced-choice task. They compared illustrations depicting three people with a) two vs. three relations, and b) three vs. six relations. Examples are in figures 3 and 4 respectively. The drawings were accompanied by a reciprocal sentence (e.g. *Danny, Guy and Omer are combing each other*) and subjects were again asked which of the two situations better describes the sentence.

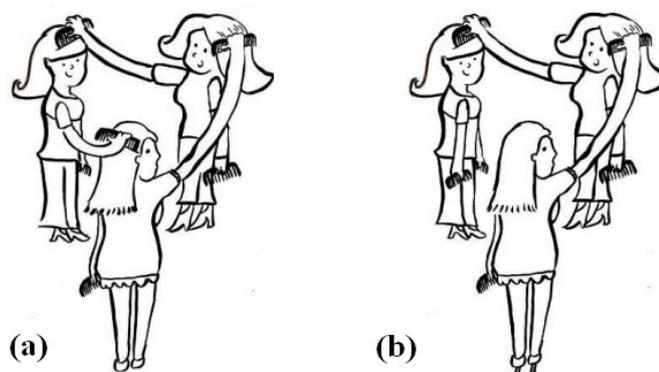


Figure 3 Interpretation of reciprocals: example of visual forced-choice method (2 vs. 3 relations) for *to comb* (KFW, 2009)

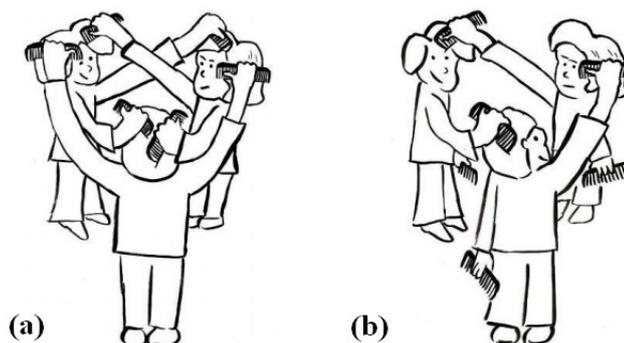


Figure 4 Interpretation of reciprocals: example of visual forced-choice method (3 vs. 6 relations) for *to comb* (KFW, 2009)

The textual experiment was an inference task. Participants were given reciprocal sentences with a subject referring to three people. They were then asked whether it is necessary or not to conclude from the sentence that one of the persons acted on another one (KFW, 2009, p.13). An example is in (12).

(12) Example of textual inference task for *to stab* (KFW, 2009)

baSana	Se'avra	omer, boaz	ve-gai	dakru	exad-et-haSeni
in-the-year	that-passed	Omer, Boaz and Guy		stabbed	one another

“Last year, Omer, Boaz and Guy stabbed one another”

ha'im	nitan	lehasik	mi-kax	she-gai	dakar	et	boaz
is-it	possible	deduce	from-this	that-Guy	stabbed	ACC	Boaz

“Can you deduce from this that Guy stabbed Boaz?”

The results of the typicality experiments clearly showed differences between predicates with respect to patient cardinality: for some predicates people preferred one patient, for some they preferred two patients, and there were also predicates that showed no preference. In the experiments testing interpretations of reciprocal sentences, KFW found 1) that three relations are always preferred over two, and 2) that three relations are chosen more often than six relations. The second result cannot be explained by the SMH nor the MIH, since they predict that the interpretation is the maximally possible one. This would indicate an interpretation involving 6 relations for most if not all of the tested predicates. Besides these separate results, the MTH predicts a positive correlation between the preference for one patient (over two) and the acceptance of a weaker reciprocal interpretation (over the strongest possible). The most important result is that the predicted relation between them was borne out (KFW, 2009). The

authors conclude that typicality effects for a predicate affect the meaning of a reciprocal sentence containing that predicate.

Chapter 3

Geometry-sensitivity

So far we have discussed three accounts that aim to explain the variety of interpretations of reciprocal sentences. According to Dalrymple et al.'s (1998) SMH, there are six logically possible meanings for the reciprocal. Each reciprocal expression receives the strongest possible meaning that is consistent with the antecedent, scope and the context in which it is uttered. One of the weak aspects of this theory is that it does not give any independent motivation for assuming the six meanings. One might ask whether these are the correct ones, and if so, whether they are the only ones. Another weak aspect is that the term 'context' is not defined explicitly yet it appears to be crucial for the theory. Once we have a theory of context, we have a clearer theory of reciprocals.

Sabato (2006) takes the SMH one step further by indeed getting rid of the vague notion of context and instead introducing the notion 'semantic restriction'. According to her version of the SMH, the interpretation of a reciprocal is derived directly from restrictions on the predicate in its scope. As a consequence, she does not need to assume a set of quantifiers. Instead, restrictions on a predicate provide us with a set of possible interpretations (graphs).

However, as we saw in our discussion of the MTH, both the SMH and the MIH seem to have a serious fault. By ignoring typicality effects, they miss out on a core aspect of the semantics of predicates. KFW (2009) point out that we do not always get a maximally possible interpretation for reciprocal sentences, as both Dalrymple et al. and Sabato would predict. According to the MTH, a reciprocal expression receives the maximally typical interpretation. This hypothesis is supported by experimental data on the relationship between the typical patient cardinality for binary predicate concepts and the interpretation of reciprocal sentences (KFW, 2009).

In this chapter, we suggest a rigorous refinement that is applicable to any one of the accounts that we discussed. We hypothesize that the interpretation of a reciprocal sentence depends on the spatial configuration of the antecedent set, which we will call the 'geometry' factor.

3.1 Circular/linear interpretations and a problem concerning typicality

We will start by having a look at a specific type of data, namely data on predicates with both linear and circular interpretations (briefly mentioned by Dalrymple et al. (p.170)). This type of predicates led us towards geometrical parameters, which appear to be relevant when dealing with interpretations of reciprocals. Consider example (13).

- (13) John, Mary and Sue are sitting alongside each other

The reciprocal sentence in (13) is generally considered to be true in a linear configuration. This means that the sentence can felicitously be uttered in case the three mentioned people are sitting in a line next to each other. Even though this appears to be an obvious fact, it is quite an interesting observation since this situation is not the maximally possible interpretation. The possibility for the three people to sit in a circle, creating more “sit alongside-relations” among the three individuals, apparently does not block the weaker, non-circular situation.

Dalrymple et al. (1998) rely on context to analyze such sentences. According to the authors, it is non-linguistic context that allows the sentence to be true for a linear configuration. For example, we would get this linear interpretation in case John, Mary and Sue are sitting on a bench. A circular interpretation can also be the correct one, but only in case the context requires it, for example when they are sitting around a campfire (p.170). Note that we are again dealing with a not clearly defined notion of context. In Sabato’s theory, which does away with context entirely, this sort of sentence clearly becomes problematic. Sabato’s MIH predicts the sentence to simply be false in the linear situation since it is possible to add one more pair of non-identical members - by closing the circle. What we observe is a non-maximal interpretation (linear) that is accepted for the reciprocal. Initially, one would suspect that we are dealing with the MTH at work in a new domain. The MTH would predict that in case the linear interpretation (which is non-maximal) is accepted for a reciprocal sentence, then this is so because it is somehow more typical than the circular one. In order to test this prediction, we would need to investigate whether a line is more typical than a circle for a predicate concept like *sit alongside*.

However, two problems arise. First of all, intuitively, it does not seem likely that linear configurations are most typical for such predicates. We believe that the different

interpretations (line or circle) are largely, if not fully, determined by contextual factors. Whether people sit alongside each other in a line or in a circle depends on factors like the surface that they sit on, the space that is available, etc. There does not seem to be an independent preference for either one of the two and thus one might wonder whether typicality has anything to do with this issue.

Moreover, even if we ignore these intuitive doubts about typicality as distinguishing such contexts, it would be hard to check such a hypothesis experimentally. Note that in KFW's experiments, typicality was treated as a phenomenon that is independent of reciprocity and thus tested separately. It was measured independently by testing patient cardinality preferences with respect to one agent in non-reciprocal situations. The interpretations of reciprocal sentences were tested afterwards in a different task. However, now we are dealing with a situation concerning (maximally) three agents. This means that the situations that we would have to compare in a typicality test are already reciprocal situations – making it hard to develop an independent measure for such a test.

We hypothesize that the linear and circular situations belong to a different class of reciprocal interpretations. When comparing the two situations (line vs. circle), we compare not only a difference in the number of relations between the three people (reflecting typicality of the predicate concept) but also a difference with respect to how they are situated relative to each other. Thus, we assume that the spatial relations between agents (their “geometry”) is a factor that needs to be taken into account in a theory of reciprocals.

3.2 A geometry-sensitive hypothesis

Let us start by looking at a specific case where this geometry-sensitivity intuition differs from the hypotheses that have so far been discussed. Consider the schematically depicted situations in figure 5.

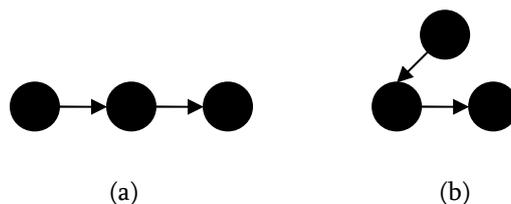


Figure 5 Two situations (schematically)

In figure 5, dots depict individuals and arrows stand for the relations between them. Note that in both 5a and 5b there are three individuals and two relations. The only difference between the two situations is the spatial configuration. In 5a, the three individuals are situated in a line whereas in 5b they are situated in a circle. Proposals like the MTH and the MIH consider these two situations equal, since they are set-theoretically equivalent once you look only at the number of individuals and number of relations. This means that for any given reciprocal sentence, for example *John, Mary and Sue are pinching each other*, the situations in figure 5 are expected to be equally acceptable (or equally unacceptable).

Our intuition is that they are not equivalent. Based on the discussion in section 3.1, we hypothesize that looking merely at the number of relations is not enough for capturing all the facts. We propose a new hypothesis that takes into account the additional factor ‘geometry’, and consequently that any mechanism that deals with reciprocals should select the reciprocal interpretation not only for a given number of individuals, but also given a certain spatial configuration. Thus, what is the maximal interpretation according to any theory is assumed to always be sensitive to geometry. According to a hypothesis that is sensitive to the spatial configuration, the acceptability rate of the two situations in figure 5 may differ. For such a hypothesis, we first consider the three individuals in a given configuration, see figure 6a and 6b. As a next step, we apply a mechanism that picks out the maximal or maximally typical meaning given that configuration.

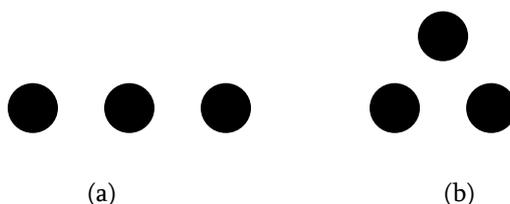


Figure 6 Two configurations (schematically)

Take for example a reciprocal sentence containing the verb *pinch*, for which it has been shown that one patient per agent is more typical than two. When we assume for example a geometry-sensitive MTH, the maximally typical interpretation for such a reciprocal sentence in configuration 6b would be such that everybody pinches one other person – thus a closed circle with 3 relations. By contrast, for figure 6a the maximally typical interpretation is expected to be one with only 2 relations. This difference indicates that we may expect the situation depicted in figure 5a to have a higher acceptability rate than 5b, since 5a is assumed to be maximally typical given the configuration while 5b is not. In 5a it is simply impossible

for the person far right to act on the person far left, while in 5b it is possible to add another relation by closing the circle. (Note that the same type of difference arises when we assume a geometry-sensitive SMH or MIH.) To summarize what we learn from this example: any hypothesis that is sensitive to geometry looks not only at the number of relations, but at the number of relations relative to the spatial configuration. We can test this hypothesis by looking at examples where such a geometry-sensitive hypothesis and the previous accounts that do not take this factor into account, make different predictions. These are exactly such cases where the number of individuals and number of relations are equal, but the configuration is different.

Note that the advantage of a hypothesis that takes geometry into account is not only empirical. It would not only be preferred because it explains a potential difference between cases such as 5a and 5b, but there is also a theoretical advantage attached to it. We saw for example that Sabato's MIH considers only those graphs that are consistent with a predicate's semantic restrictions, when evaluating a reciprocal sentence. Let us call these graphs the "reference set". The reference set for a given situation is determined by properties of the predicate, and is thus a subset of the set of all possible graphs for reciprocals. A hypothesis that takes geometry to be an additional factor, on top of the properties of the predicate, restricts the reference set even further. This means that when evaluating a reciprocal sentence, we only consider the alternatives that are consistent with a) the geometrical configuration and b) the predicate. This makes such a hypothesis cognitively even simpler and thus preferred. The difference between reference sets for a geometry-sensitive hypothesis and a non-geometry-sensitive hypothesis is clarified schematically in figures 7a and 7b. For reasons of simplicity, we do not worry about the direction of the relations in this figure.

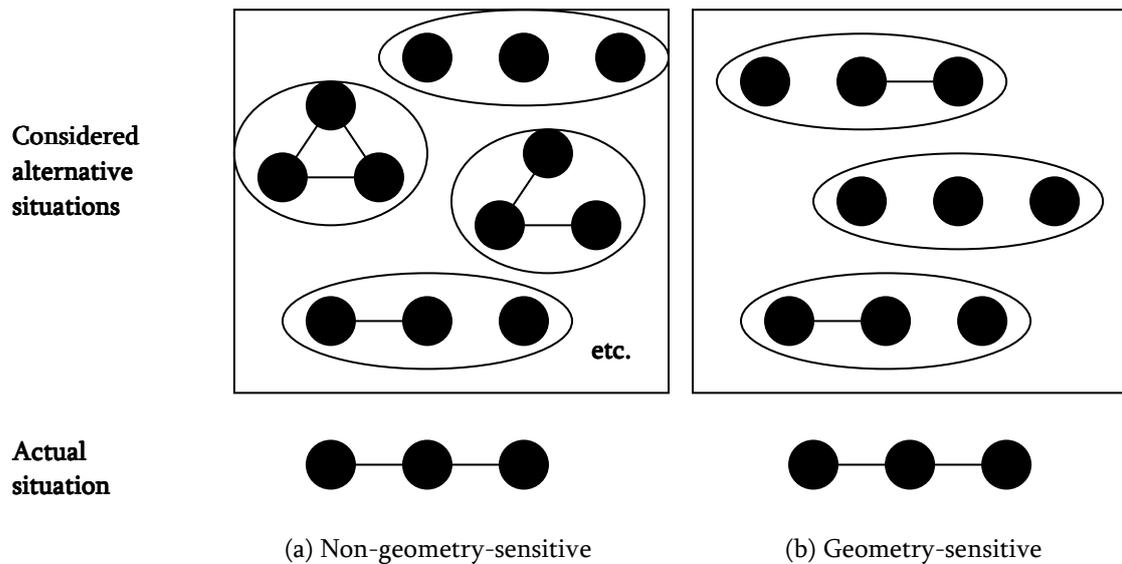


Figure 7 Possible reference sets according to non-geometry-sensitive vs. geometry-sensitive accounts

The box in figure 7a is an example of a possible reference set for the given situation according to an account that does not care about geometry. It includes all possible interpretations that contain three individuals (and are consistent with the predicate’s restrictions). Figure 7b gives a possible reference set for the same situation but taking into account the geometrical configuration of the situation that is being evaluated. Thus, this set only contains possibilities where the three individuals are situated in a line.

3.3 Subconclusion

In chapter 2 we have seen that it is quite a difficult task to come up with a theory that explains all the variation in interpretations of reciprocal expressions. We have gone over two types of them: 1) two accounts that claim that a reciprocal sentence always receives the maximally possible interpretation, looking only at definitional aspects of meaning, and 2) an account that claims that instead it receives the maximally typical interpretation, thus looking also at typicality effects for binary predicate concepts. The three proposals combined have provided us with a vast amount of insight into the issue of reciprocals and have given us the opportunity to build on them in order to account for more data. The observations and intuitions concerning predicates with both linear and circular interpretations that have been presented (both in chapter 3 of the current thesis and briefly in Dalrymple et al. (1998)), led us towards a

refinement of the theories so far. We proposed a geometry-sensitive hypothesis, which takes the discussed accounts as its starting point but adds the factor ‘geometry’ as a crucial notion in determining what is maximal for a reciprocal sentence.

At this point, our ideas are merely a hypothesis, which is based mainly on intuitions. Chapter 4 reports on an experimental study that was conducted in order to verify whether geometry indeed plays a role when interpreting a reciprocal. The experiment deals with cases where the SMH, MIH or MTH on the one hand make different predictions from a geometry-sensitive hypothesis on the other hand.

Chapter 4

Experimental study

An experimental study has been conducted in order to test the geometry-sensitive hypothesis, according to which “geometry” is a factor in determining the interpretation of reciprocal sentences. The main issue under investigation is whether there can be a difference in acceptability between two set-theoretically equivalent situations that differ solely in geometric configuration. In this experiment, we aimed to answer the following specific research question: *Is there a difference in acceptability rates for reciprocal sentences in linear vs. open circle situations?* The two situations that were tested contain three individuals, and are depicted schematically in figure 8 below.

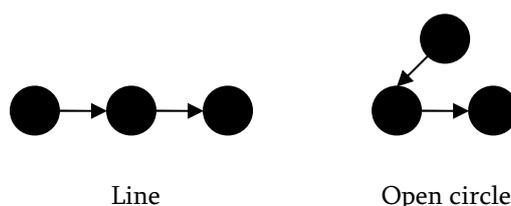


Figure 8 The two situations that will be compared in the experiment (schematically)

In order to compare acceptability for the two configurations, participants were given an acceptability judgement task in which they were asked to indicate whether they accept a reciprocal sentence in a given situation or not.

Section 4.1 reports three pilot studies that were conducted prior to the actual experiment. Based on results from and comments on these studies, we have modified both methods and materials and eventually conducted the experiment that is discussed in detail in section 4.2.

4.1 Pilot studies

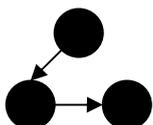
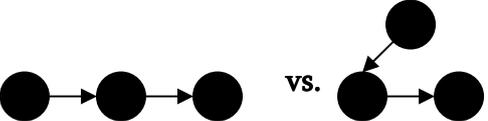
Before running a large scale experiment, pilot studies were conducted that were aimed at testing whether the measures that we intended to use actually behave as we theoretically expected them to.

Pilot study 1: Schematically (verbal)

The first pilot study consisted of two tasks: an acceptability judgement task and a forced choice task (in Dutch). The acceptability judgement task measured truth values per configuration (line and open circle), while the forced choice task checked which configuration (line or open circle) was preferred. A total number of 5 subjects (2 male, 3 female) participated in this study. Each subject took part in a one-to-one verbal session with the experimenter.

Prior to the tasks, subjects were given a brief introduction explaining the task. For each item, the experimenter described a situation concerning three people and drew that situation schematically (comparable to figure 8 above). Subjects were then given a sentence and were asked for their personal judgement on whether the described situation is a possible depiction of that sentence. The test items for the acceptability judgement task were lines and open circles for three verbs: *lean on*, *pinch* and *comb* – resulting in 6 acceptability judgement questions for each subject. The sentences that accompanied the drawings were reciprocal sentences, for example *A, B and C are leaning on each other* (Dutch: *A, B en C zijn op elkaar aan het leunen*). In case a subject accepted both the line and the open circle for a particular verb (2 x “yes”), a forced choice question followed, asking for the preferred configuration for the sentence. The same was done in case a participant accepted neither of the two configurations for a given verb (2 x “no”). The third and last possibility is that subjects accepted only one of the two configurations (1 x “yes” and 1 x “no”). In that case, the preference task for that verb was redundant and we counted the one configuration that they accepted as the preferred configuration. A summary of the test items for pilot 1 is given in table 2.

Table 2 Summary of test items in pilot study 1

<i>Verbs</i>	<i>Task</i>	<i>Condition</i>	<i>Situation</i>	<i>Sentence</i>
Lean on, Comb, Pinch	Acceptability	-		A, B and C are <i>X</i> -ing each other
	Acceptability	-		
	Preference	Only if acceptability task results: - 2 x “yes” or - 2 x “no”		

The results of pilot study 1 are summarized in table 3. The linear configuration was accepted 40% of the time while the open circle was accepted only 13% of the time. Also, no subject accepted the circular configuration for a verb without accepting the linear one for that same verb. Finally, subjects preferred the line over the open circle 92% of the time⁶.

Table 3 Pilot study 1 results

	Acceptability <i>“Yes” response (%)</i>	Acceptability <i>“No” response (%)</i>	Preference <i>(%)</i>
Line	40	60	92
Open circle	13	87	8

Thus we see that in this pilot there is a difference in acceptability of reciprocal sentences in two situations that differ merely in configuration: the linear configuration appears to be accepted more often than the open circle configuration. These results have given us a first indication that the interpretation of reciprocal sentences might indeed be sensitive to geometry.

Pilot study 2: Picture-based (verbal)

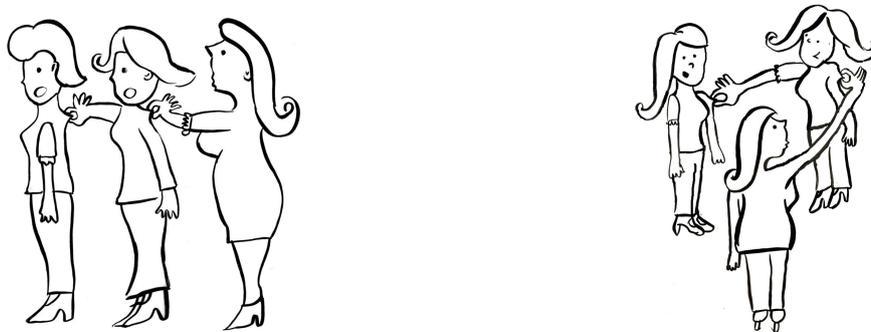
Pilot study 1 only made use of schematic drawings. The second pilot study aimed to see whether the same results are obtained once we deal with situations concerning people. It was comparable to the first pilot study, but used pictures instead of merely schematic drawings consisting of dots and arrows. Also, a larger number of verbs was tested.

This pilot again consisted of two tasks: an acceptability judgement task and a forced choice preference task (in Dutch). A total number of 4 subjects participated (2 male, 2 female) and each subject took part in a one-to-one verbal session with the experimenter. The experimenter explained the tasks in an introduction. For each question, subjects were shown a card with a drawing containing three people, then they were given a sentence and they were asked whether the described situation was a possible depiction of that sentence according to them. Each subject was shown two drawings (line and open circle) for a number of verbs (8 or 16)⁷. The two drawings per verb differed solely with respect to the configuration of the three

⁶ Preference was calculated either based on the preference task (directly), or based on a difference in acceptability between the two configurations in the acceptability task (indirectly). In case only one of the two configurations was accepted, this configuration was considered the preferred one.

⁷ Two subjects were tested on 16 verbs while the other two subjects were tested on a subset of these verbs 8 verbs. This was due to different amounts of time that each subject had available for the study.

individuals in it. The eight verbs that all four participants were tested on are *stab*, *scrape*, *clean*, *comb*, *paint*, *point at*, *push* and *touch*. The additional eight verbs that were given to two of the four participants are *apply make-up*, *wipe*, *shake*, *pinch*, *hit*, *lean on*, *tickle* and *pat*. The order of configuration was counterbalanced: for half of the verbs the line was shown first, for the other half the open circle was shown first. The sentences given were reciprocal sentences, for example: “*A, B and C are pinching each other*” (Dutch: *A, B en C zijn elkaar aan het knijpen*). Similar to pilot study 1, an explicit preference task followed only in case a subject accepted both configurations for a given verb or refuted both. In case only one of the configurations was accepted, this configuration was counted indirectly as the preferred one. An example of a set of test items is given in figure 9. At the end of each session, there was a brief discussion with the subject on the clarity of drawings and task.



Question: Do you think this drawing is a possible depiction of the sentence *A, B and C are pinching each other?*

Question: Do you think this drawing is a possible depiction of the sentence *A, B and C are pinching each other?*

Figure 9 Examples of acceptability judgement task (2 items) for pilot study 2

The results of pilot study 2 are summarized in table 4 below. The linear configuration was accepted 83% of the time while the open circle was accepted 44% of the time. Moreover, the line was preferred over the open circle 71% of the time⁸.

Table 4 Pilot study 2 results

	Acceptability “Yes” response (%)	Acceptability “No” response (%)	Preference (%)
Line	83	17	71
Open circle	44	46	29

⁸ Preference was again calculated either based on the preference task (directly), or based on a difference in acceptability between the two configurations in the acceptability task (indirectly).

These results show the same pattern as pilot study 1, though the differences between the line and open circle are slightly smaller. Based on post-experiment discussions with subjects, a few changes were made in the drawings and an important change was made in the formulation of the sentences in the task. The Dutch sentence that was initially given to subjects was “*A, B en C zijn elkaar aan het X-en*”. This type of formulation led subjects to interpret the sentence as describing an ongoing event. They reported that they often accepted the open circle configuration because they interpreted it as a snapshot of a larger event, one that continued before and after this snapshot. In such a larger event, all three people in the drawing were acting on a patient – creating a closed circle. Based on these remarks, we changed the formulation to “*In deze tekening X-en A, B en C elkaar*”. We tested this new formulation with both the participants that explicitly reported having interpreted the tested sentences as describing an ongoing event, as well as with other participants. We found that the slightly different formulation of the reciprocal sentence combined with the addition of “in this drawing” causes a more static interpretation and forces subjects to interpret the situation depicted instead of a situation that is merely suggested.

Pilot study 3: Picture-based (written)

The third pilot study consisted of short written questionnaires of 15 questions (including fillers) that had the same structure as the questionnaires of the intended experiment. Ten subjects (1 male, 9 female) participated. The first part of the questionnaire contained an acceptability judgement task (line and open circle configuration), the second part contained a preference task (line vs. open circle configuration). The questionnaires that we used contain similar drawings to the ones that were used in pilot study 2.

There were two versions of the questionnaire, testing a total number of 6 verbs (*lean on, comb, point at, pull, shake* and *stab*). Each version tested 3 of these verbs in the acceptability task and the other 3 verbs in the preference task. Fillers were equal for version 1 and 2. For the acceptability part of the questionnaire, the fillers were drawings depicting three people with a) only one relation between them (one relation), b) three relations between them (closed circle) or c) all of the people acting on themselves (act on self). The fillers for the preference part of the questionnaire were either a) two open circle drawings, but with females vs. males, or b) a line vs. a closed circle. A summary of all items in the questionnaires is in table 5.

Table 5 Summary of test items in pilot study 3 (OC = Open circle, M = Male, F = Female)

<i>Version 1</i>	<i>Version 2</i>	<i>Fillers (same for both versions)</i>
Part 1: Acceptability (6)	Part 1: Acceptability (6)	Acceptability fillers (4)
Lean on line & OC Comb OC & line Point at line & OC	Pull OC & line Shake line & OC Stab OC & line	- Paint one relation - Push one relation - Wipe act on self - Scrape closed circle
Part 2: Preference (3)	Part 2: Preference (3)	Preference fillers (2)
Pull OC vs. line Shake line vs. OC Stab OC vs. line	Lean on line vs. OC Comb OC vs. line Point at line vs. OC	- Point at OC M vs. F - Scrape line vs. closed circle

The order of configuration was counterbalanced over verbs: for half of the verbs the line occurred first in the questionnaire, for the other half of the verbs the open circle occurred first. Otherwise items for both parts of the questionnaire were randomly ordered. The exact order of questions for both versions is in appendix A1. An example of a complete questionnaire including instruction page is in appendix A2. Subjects were randomly divided over versions.

The overall results of pilot study 3 are in table 6. The acceptability task shows roughly the same overall pattern as pilot study 1 and 2, and supports a geometry-sensitive hypothesis.

Table 6 Pilot study 3 results (total)

	Acceptability <i>"Yes" response (%)</i>	Acceptability <i>"No" response (%)</i>	Preference <i>(%)</i>
Line	63	37	47
Open circle	33	67	53

The line configuration was accepted in 63% of the cases, while the open circle was accepted in only 33% of the cases. Interestingly, the preference task appears to show a guess pattern (47% vs. 53%). This is not what one would expect assuming a hypothesis that is sensitive to the geometrical configuration, since such a hypothesis predicts a clear preference for the line. Table 7 gives a more detailed view of this result, by looking at preference results per verb.

Table 7 Pilot study 3 results (per verb)

	Preference (%)					
	Lean on	Comb	Point at	Pull	Shake	Stab
Line	60	40	20	40	80	40
Open circle	40	60	80	60	20	60

Table 7 shows that the results for the verbs are mixed. The only verb that shows a clear preference for a linear configuration is *shake*. The verb *point at* shows the opposite pattern, while the results for the remaining four verbs are a lot less clear.

What we see is a discrepancy between the results for the acceptability task (the line is accepted more often than the open circle) and the preference tasks (there is no preference). Due to these results, it is necessary to critically evaluate the two measures that were used. In the acceptability judgement task, subjects were asked to answer “yes” or “no” to the question: *Do you think this drawing is a possible depiction of the sentence?* This task unambiguously measured the truth value for each reciprocal sentence given a situation. Once subjects are being asked about their preference, it is much less clear what we are measuring theoretically speaking. A possible explanation for the discrepancy in results between the two tasks in this final pilot, is that subjects used different strategies for both tasks, since it is not clear what it means “to prefer” a situation once we are dealing with reciprocal sentences. Thus, it is also not clear how to interpret results of such a preference task. Based on the questionable results and the fact that we do have a clear task measuring truth values, we decided to leave the preference task out of the actual experiment.

4.2 The experiment

The experiment consisted of an acceptability judgement task in the form of a written questionnaire. It measured the acceptability of reciprocal sentences in Dutch in two minimally different situations containing three individuals: line versus open circle. If we find a difference in acceptability rates between two different configurations with an equal amount of relations, then we can conclude that the interpretation of reciprocal sentences is sensitive to the geometrical configuration of individuals.

Method

Participants We tested two groups of undergraduate students at Utrecht University. The total number of subjects is 71, of which 3 were excluded due to a native language other than Dutch. Analyses were done for 68 native speakers of Dutch (20 male), mean age 22.34. All participants were students at the Faculty of Humanities, with most participants coming from either the department of Linguistics (21%), Cognitive Artificial Intelligence (26%), Speech and

Language Pathology (21%) or Liberal Arts and Sciences (10%). Participants were recruited via instructors of undergraduate courses and they voluntarily participated in the study under consent of their instructors.

Materials Materials consisted of questionnaires containing 33 questions: 22 test items and 11 filler items. Each page of the questionnaire contained two questions, resulting in a questionnaire of 17 pages long. We created two questionnaires (A and B), with each one testing 11 verbs. Every participant received one of the two questionnaires.

The experiment contained 14 verbs that were also tested by Kerem, Friedmann and Winter (2009) in a picture-based experiment with a different research question. In addition, we added 8 new verbs - resulting in a total number of 22 verbs. An overview of all verbs per questionnaire is in table 8.

Table 8 Lists of verbs tested in the experiment (per questionnaire)

	<i>Questionnaire A</i>	<i>Questionnaire B</i>
1	Stab	Push
2	Scrape	Lean on
3	Apply make up	Pull
4	Wipe	Touch
5	Clean	Tickle
6	Comb	Pat on the back
7	Shake	Kick
8	Pinch	Hold [hands]
9	Paint	Shoot
10	Point at	Play [sound]
11	Hit	Wet

For each verb, a questionnaire contained two test items: one for a line configuration and one for an open circle configuration, resulting in 22 test items in total per questionnaire. All test items contained a picture with three individuals and two relations between them. The two items for each verb differed only with respect to configuration. An example of a pair of test items is in figure 10.

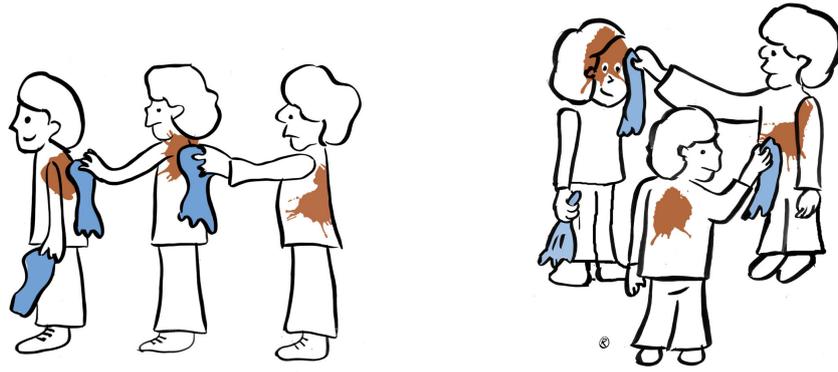


Figure 10 Example pair of test items (*to clean*) in a line and an open circle configuration

Apart from the test items, each questionnaire contained 11 filler items. The filler drawings depicted the same 11 verbs as the test items, but with a difference in number of individuals, number of relations or with incorrect verbs in the accompanying sentence. Part of the filler drawings was clearly acceptable for a reciprocal sentence, namely those with three people and three relations (closed circle). There were also fillers that were clearly not acceptable, namely those with three people and only one relation. Moreover, the questionnaires contained fillers with drawings containing ten people in a line and 9 relations (long line) and ten people in a circle and 9 relations (big open circle). Finally, there were fillers that were purely distracters (three people acting on themselves, sentences with an incorrect verb).

Items were pseudo-randomly ordered, with several restrictions. Since there were three items per verb (2 test items and 1 filler), we made sure that these items did not influence each other. First of all, the mean distance between items containing the same verb was 10.82 questions. The minimal distance was 5 questions, so that items depicting the same verb were never on the same page nor on pages immediately following each other. Secondly, a closed circle filler never preceded the line and open circle test items for that same verb, so that it could not influence answers to test items. Thirdly, the ordering of line and open circle test items was counterbalanced: for half of the verbs in each questionnaire the line preceded the open circle while for the other half the open circle preceded the line. Finally, we used two versions of each questionnaire, where filler items remained in place but test items for each verb were switched so that the line and open circle occurred in the opposite order from the other questionnaire, accounting for a potential ordering effect. The complete structure of all four questionnaires can be found in appendix B1.

Task We made use of an acceptability judgement task. Each question consisted of a drawing and a reciprocal sentence. In the instruction of the questionnaire, subjects were informed that we were interested in their opinion. They were asked to judge for each item whether the depicted situation in the drawing was a possible depiction of the given sentence. The given options were “yes” and “no” and subjects were expected to circle the answer of their choice. An example of a question as it appeared in a questionnaire is in figure 11. A complete questionnaire including instruction can be found in appendix B2.

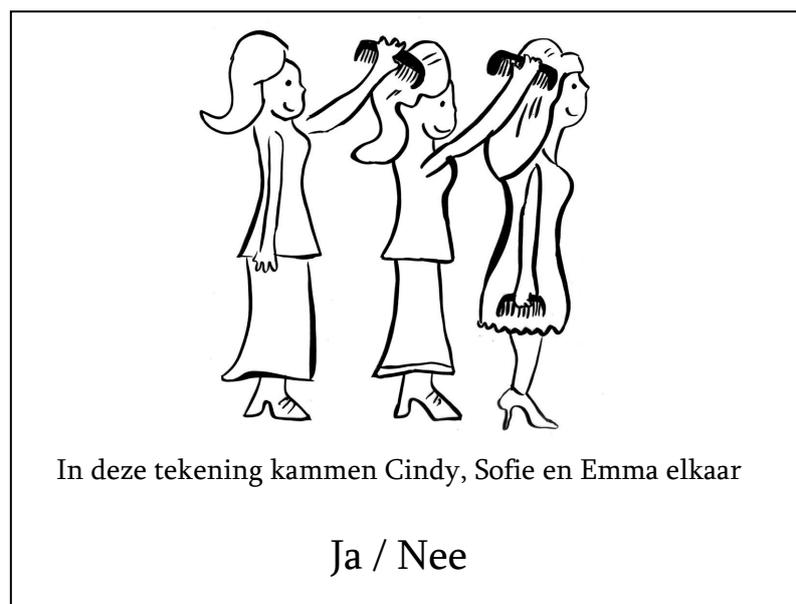


Figure 11 Example of a question (*to comb*, line configuration)

Procedure Participants were tested per group for approximately 10 minutes in a large classroom, either at the start of a class or during a break. The experimenter instructed participants to read the instruction and then wait until all participants were done reading. They then received a signal, which indicated that that they were allowed to turn the page and start the questionnaire. The total amount of time given for the questionnaire (10 minutes) was based on a test-run with 2 participants prior to the actual experiment. This amount of time proved enough for all participants to finish the questionnaire.

Data analysis The design of the experiment is a 2 (Configuration) x 2 (Verb group) x 2 (Order) x 2 (Participant group) mixed factorial design. The main independent variable Configuration (line vs. open circle) is a within-subjects factor (or repeated-measures variable), since each subject judged both the line and the open circle configuration for a given verb. The

remaining three variables are between-subjects factors: Verb group (A vs. B⁹), Order (1 vs. 2) and Participant group (1 vs. 2). Acceptability (in proportions) is the dependent variable, calculated from the proportion of “yes” answers. The data were analyzed using two separate 2 x 2 x 2 x 2 repeated measures Analysis of Variance (ANOVA): one over subjects and one over items. In the ANOVA over items, the within-items factors are Configuration, Participant group and Order and the between-items factor is Verb group.

Results

Acceptability (over subjects) The data that were analyzed in the first repeated measures ANOVA are mean proportions per participant. The ANOVA compares the mean proportion of “yes” replies for the line configurations to the mean proportion of “yes” replies for the open circle configurations and tests the null hypothesis that both means are equal. The alternative hypothesis (H₁) is that the means differ from each other.

We ran a mixed ANOVA with Configuration as the within-subjects factor and Verb group, Order and Participant group as the between-subjects factors. Table 9 shows a descriptive overview of the results. It contains the mean proportions per verb group per order.

Table 9 Acceptability for line and open circle: mean proportion (standard deviation)

	Group 1				Group 2			
	Verb group A		Verb group B		Verb group A		Verb group B	
	Order 1	Order 2	Order 1	Order 2	Order 2	Order 2	Order 1	Order 2
Line	.22 (.39)	.19 (.35)	.18 (.15)	.32 (.26)	.25 (.32)	.11 (.28)	.18 (.30)	.17 (.20)
Open circle	.24 (.41)	.25 (.35)	.19 (.20)	.27 (.26)	.22 (.38)	.09 (.29)	.17 (.30)	.15 (.20)

The main effect of configuration was not significant ($F(1,60) = .068; p = .796$). This means that we did not find a difference in the acceptability of reciprocal sentences in linear vs. open circle configurations. There were no significant between-subject effects (all $F_s < 1$) and none of the two-way or three-way interactions were significant (all $F_s < 1$).

Acceptability (over items) For the second repeated measures ANOVA, we used mean proportions per verb. The ANOVA compares the mean proportion of “yes” answers for the line configuration to the mean proportion of “yes” answers for the open circle configuration and again tests the H₀ that these means are equal (with the H₁ that the means are not equal).

⁹ See table 8 for an overview of all verbs in groups A and B.

In the second mixed ANOVA, the within-items factors were Configuration, Participant group and Order. The between-items factor was Verb group. Similar to the first analysis, we found no main effect of configuration ($F(1,20) = .054$; $p = .818$). There was however a significant main effect of participant group ($F(1,20) = 31.279$; $p < 0.001$). This means that in the item analysis, the two groups of participants differed from each other in their replies. The order of presenting the verbs had no effect ($F(1,20) = .547$; $p = .468$). We also found that there is no between-items effect of verb group on replies ($F(1,20) = .039$; $p = .846$). This means that we did not find a difference in mean acceptability rates of reciprocal sentences between the two different groups of verbs.

There were two significant interaction effects. We found an interaction between Verb group x Order of presenting the verbs ($F(1, 20) = 13.148$; $p = .002$) and an interaction between Participant group x Order of presenting the verbs ($F(1,20) = 31.079$; $p < .001$). Regarding the first interaction effect (Verb group x Order), a pairwise comparisons analysis¹⁰ revealed that for verb group A, the overall acceptability in questionnaires with order 1 is higher than in those with order 2 ($t(20) = 3.13$, $p = .006$) while verb group B shows a trend in the opposite direction ($t(20) = 2.09$, $p = .055$). For the second interaction effect (Participant group x Order), we found that for participant group 1, overall acceptability is higher in questionnaires with order 2 than in those with order 1 ($t(20) = 2.27$, $p = .035$). Participant group 2 shows the exact opposite: acceptability in questionnaires with order 1 is higher than in those with order 2 ($t(20) = 4.41$, $p < .001$).

The main effect of participant group and in particular the two interaction effects that the item analysis showed, are not easily explainable. However, they indicate that we should not look merely at total means. Therefore we also looked at the data on the individual verbs. Table 10 gives an overview of mean acceptability rates for line and open circle per verb.

Table 10 Acceptability for line and open circle per verb: mean proportion

	<i>Verb</i>	<i>Line</i>	<i>Open circle</i>
1	Stab	.15	.24
2	Scrape	.15	.24
3	Apply make-up	.21	.15
4	Wipe	.21	.24
5	Clean	.18	.21
6	Comb	.24	.18
7	Shake	.21	.21

¹⁰ For the pairwise comparisons analysis, we used the Bonferroni method with significance levels that are adjusted for multiple comparisons.

8	Pinch	.18	.15
9	Paint	.26	.26
10	Point at	.09	.06
11	Hit	.21	.24
12	Push	.09	.18
13	Lean on	.29	.44
14	Pull	.12	.12
15	Touch	.32	.18
16	Tickle	.15	.21
17	Pat on the back	.12	.12
18	Kick	.09	.18
19	Hold [hands]	.71	.44
20	Shoot	.15	.09
21	Play [music]	.12	.09
22	Wet	.15	.09

Next, figures 12 and 13 are box plots which indicate whether there are any outliers among the verbs. Figure 12 shows this for the acceptability rates per participant group per order (11 = Participant group 1, Order 1; 12 = Participant group 1, Order 2; etc.). Figure 13 shows the outliers in overall acceptability rates for line and open circle.

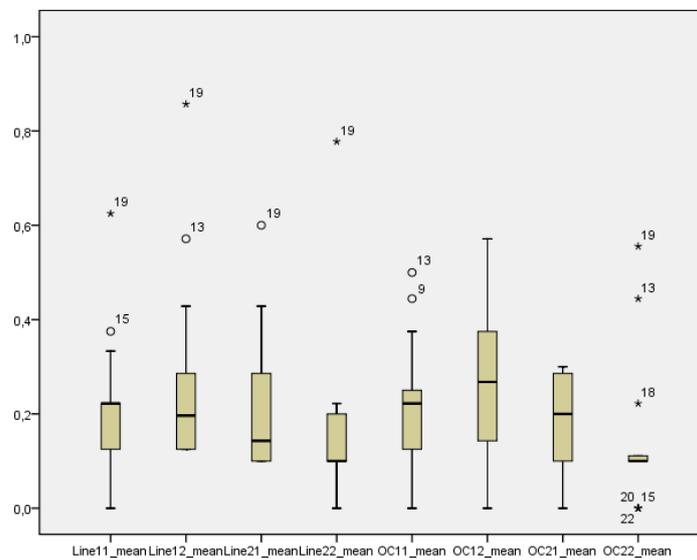


Figure 12 Outliers (per participant group per order)

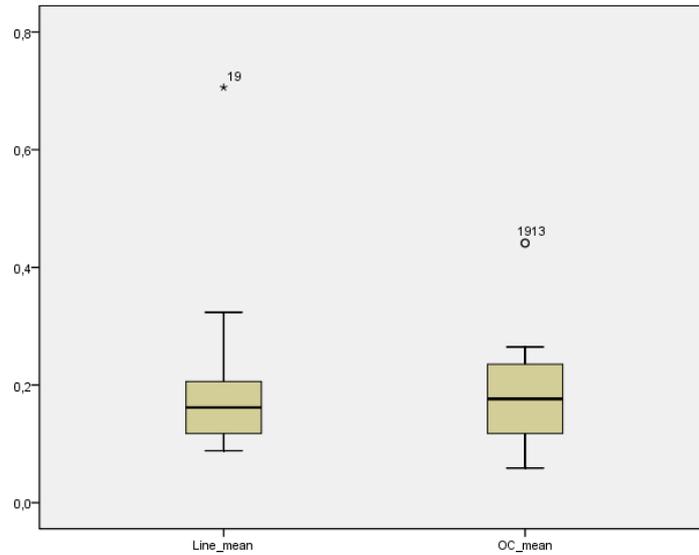


Figure 13 Outliers (overall)

In figure 12 we see that verb number 19 (*hold hands*) is an outlier in five out of eight possible combinations of participant group and order. Verb number 13 (*lean on*) occurs three times. When we collapse these data into overall outliers, verb 19 remains a clear outlier (especially for the line configuration) and verb 13 is also still present. In table 10 we see that for *hold hands* it is the case that the line was accepted 71% of the time while the open circle was accepted only 44% of the time. This difference in proportion between line and open circle is significant ($N = 34$, exact $p = .012$). Thus the linear configuration was accepted significantly more often than the open circle configuration for the verb *hold hands*. Another verb that shows the same pattern is *touch* (verb 15), but the difference in acceptability for that verb fails to reach significance ($N = 34$, exact $p = .180$). Verb 13 (*lean on*) shows the opposite pattern, but the difference is not significant ($N = 34$, exact $p = .267$).

Chapter 5

Discussion

The main issue throughout this thesis has been how to account for the variety of interpretations of reciprocal sentences. More specifically, we aimed to see whether “geometry” is one of the explaining factors. We started out by observing that a sentence like *John, Mary and Sue are sitting alongside each other* is true in a linear configuration, even though previous accounts predict it to be false or are unable to give a clear explanation for why it should be true. In an experiment, we compared acceptability rates for a linear situation vs. an open circle situation in order to see whether geometry matters in case all else is held equal. The current chapter discusses the main findings of the experiment, proposes a revised hypothesis and elaborates on possibilities for further research.

Main findings In our first analysis (ANOVA over subjects) we found no significant effects at the level of participants. There was no overall difference between the acceptability rates for a line and an open circle, which is against the predictions of a geometry-sensitive hypothesis. There was also no difference between the replies for the two groups of verbs. This is not against our hypothesis, since we expected a potential effect to generalize over verbs,

The second analysis (ANOVA over items) showed different results. We again failed to find a main effect of configuration, but there were two interaction effects that were not quite straightforward. We found an interaction between verb group and the order of questions and an interaction between participant group and the order of questions. These effects were unexpected and cannot easily be explained. Moreover, one might wonder whether these effects are of any relevance. When we look at the mean acceptability rates per order, per verb group or per participant group, we see that there is a clear floor-effect across all conditions (see table 9). This means that acceptance is very low throughout conditions (around 20%) and that the interaction effects are thus based on differences around this overall very low acceptability rate. This makes it very difficult to give them a meaningful interpretation.

A closer analysis of each verb separately revealed that the verbs *hold hands* and *lean on* were outliers. For *lean on*, the open circle was accepted 44% of the time while the line was accepted only 29% of the time. However, this difference was not significant. Also, there is an

alternative explanation for why this verb is an outlier, which has to do with the open circle drawing for *lean on* (see figure 14). For this particular drawing, several participants explicitly indicated that they were not sure whether all three people in the picture were leaning or not. Also, many participants who claimed that they had only accepted reciprocal sentences in cases where each individual was an agent, still accepted the sentence accompanying this particular drawing – implying that they thought that there were three relations instead of the intended two.



Figure 14 Open circle drawing for *to lean on*

This ambiguity in the drawing might explain why *lean on* is an outlier that tends towards a preference for the open circle. More interesting is the outlier *hold hands*, for which we found a significant effect: the line was preferred over the open circle. Such an effect is predicted by a geometry-sensitive hypothesis. As mentioned in the results section, one other verb showed a trend towards this pattern (but without a statistically significant difference), namely *touch*. Comparing these two verbs (*hold hands* and *touch*) to the other verbs, we note an interesting characteristic that differentiates them. The two verbs under consideration are the only two symmetric verbs among all verbs in the questionnaires.

Thus our main results are twofold: 1) in general the line and the open circle are equally unacceptable (around 20% “yes” responses), 2) for two symmetric verbs the line seems a more plausible option, and it is accepted more often than the open circle (one significant effect and one trend). From the first main result, we can conclude that the hypothesis on geometry-sensitivity in the broad sense is not supported. Most likely, the factor “geometry” is not general in the sense that it influences acceptability of reciprocal sentences with any given verb. For the second result, there is a possible explanation which is still in line with our intuition on the role of geometry. This explanation allows us to rephrase the hypothesis in such a way that it is more specific – and again testable.

Let us take a closer look at the differences between non-symmetric and symmetric verbs. What we initially aimed to compare were two minimally different situations containing three individuals and two relations between them. These situations are indeed testable for non-symmetric verbs (like *scrape*, *pinch*, etc.). However, for the symmetric verbs *hold hands* and *touch*, we actually compared two other situations. This is clarified in table 11.

Table 11 Compared situations for non-symmetric vs. symmetric verbs

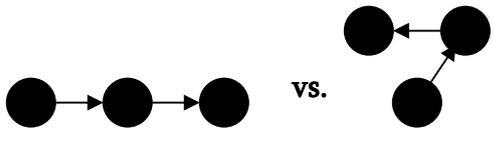
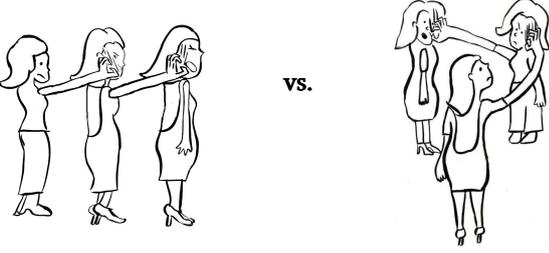
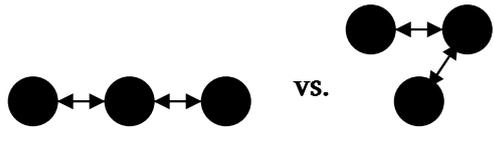
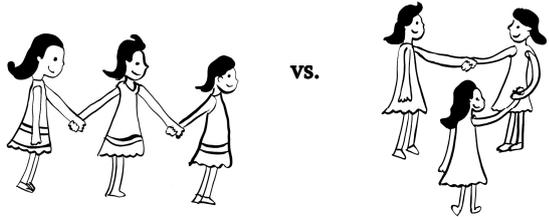
<i>Compared situations (schematically)</i>	<i>Example</i>
<p>Non-symmetric</p> 	
<p>Symmetric</p> 	

Table 11 clarifies the difference in compared situations for non-symmetric vs. symmetric verbs. For the non-symmetric verbs (for example *scrape*), we indeed compared situations with two relations among the three individuals. However, for the two symmetric verbs, we actually compared situations with four relations among the three individuals. What we found is that geometry-sensitivity appears to play a role for the symmetric verbs, whereas for the non-symmetric verbs reciprocal sentences are simply judged to be false in both of the situations.

Revised geometry-sensitive hypothesis

The explanation that we propose for these results is that geometry-sensitivity is not as strict as we initially thought. Our revised notion of geometry-sensitivity goes as follows. When evaluating a reciprocal sentence, only the position of individuals in space is determined. For example, when we evaluate a sentence in a linear situation, we do not consider alternative “geometrical shapes” (e.g. a circle). However, unlike what we previously thought, different directionality of actions may in fact be considered. This means that we consider situations where the individuals turn around into a different direction in order to act on (more) patients. Consequently, the linear situation for the non-symmetric

verbs is not maximal after all. The reference set contains situations with more relations among the three people, as is shown in figure 15b. The exact same story holds for the open circle situations.

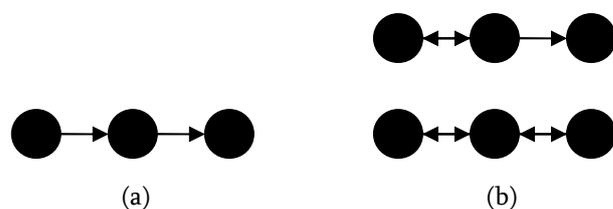


Figure 15 Actual situation (a) and considered alternatives (b)

If we assume this revision of the hypothesis, we explain why participants do not accept either of the situations for all the non-symmetric verbs. However, note that the SMH, MIH and MTH can also explain these facts. Geometry-sensitivity simply does not surface here, since both situations are considered to be false. Thus, the data on the non-symmetric verbs does not support nor refute our geometry-sensitive hypothesis in its revised formulation.

However, with this revised hypothesis we can explain why participants do appear to accept the line for symmetric verbs, and why they appear to accept it more often than the open circle. For these verbs, the situations depicted contain four relations. The modified hypothesis predicts that this number of relations is maximal given the linear configuration, since 1) the position of individuals is fixed, so we do not consider alternatives other than a linear configuration and 2) turning individuals around into a different direction then cannot lead to more relations. On the other hand, four relations is not maximal given the open circle configuration. In the open circle, it is possible to add more relations (without losing typicality) – by closing the circle.

Thus, we have seen that the result on the symmetric predicates suggests that a (revised) geometry-sensitive hypothesis holds. However, this result is quite meagre since it is based on data from merely two verbs. We do however now have a clear hypothesis that enables us to test whether the same result is obtained for other symmetric verbs. Also, we have an explanation for why the acceptability measure for *touch* did not manage to show a significant effect. Since the drawings in the questionnaires were not made with a difference between symmetric and non-symmetric verbs in mind, we did not pay attention to the clarity of them with respect to symmetry. The verb *touch* is ambiguous between being symmetric or non-symmetric and we believe that the drawings that we used to depict *touch* prime the non-symmetric reading. Consider the drawings in figure 16.

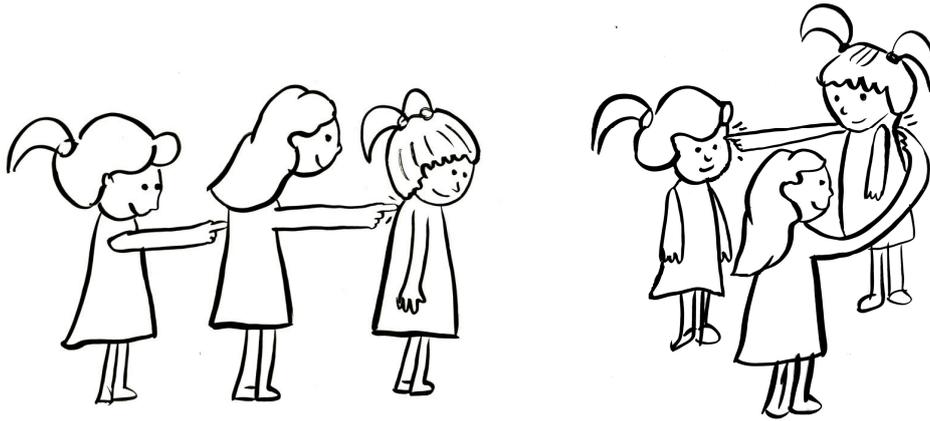


Figure 16 Line and open circle drawings for *to touch*

These drawings can very easily be understood as depicting a non-symmetric reading of the verb *touch*. This means that the girls touch intentionally, so that when girl A touches girl B this does not entail that girl B is also touching girl A. However, one can also interpret the verb *touch* symmetrically, where it means something like “being in contact”. In that case, girl A touching girl B entails also the reverse. The revised geometry-sensitive hypothesis would predict the difference in acceptability for line and open circle to be significant once we disambiguate the drawings and clearly depict a symmetric reading of *touch*.

Further research and remaining questions In order to support the revised hypothesis, additional research that tests a larger amount of (unambiguously) symmetric verbs is needed. Again, we can measure and compare acceptability rates for two minimally different situations: a line and an open circle. Since we are then dealing with symmetric verbs, we would compare situations with four relations (instead of two) among three individuals or objects, with merely a difference in configuration.

We believe that there are at least two groups of symmetric verbs that are easily testable. First of all, there is a class of verbs describing a mechanical process that results in the joining of two objects or materials. The result of the process can be photographed or drawn. Examples of such verbs are *to be welded to*, *to be soldered to*, *to clinch*, *to be pinned to*, *to be nailed to*, *to be screwed to*, *to be glued to*, *to adhere* or the more general *to be connected to*. Imagine for example a situation where there are three wooden planks and two nails (or “*to be nailed to*-relations”). With these components, we can now distinguish between a linear situation and a circular situation, as in figure 17.

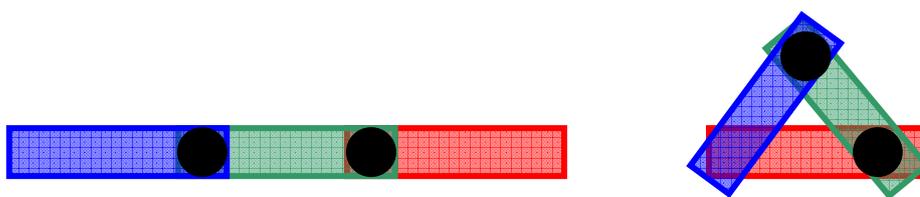


Figure 17 Line and open circle situations for *to be nailed to*

As in the experiment that was conducted for the current thesis, acceptability of reciprocal sentences can be measured for each situation. For the examples in figure 17, the sentence would be *The blue plank, the green plank and the red plank are nailed to each other.*

A second group of symmetric verbs describe some sort of contact between people, also containing the verbs *hold hands* and (symmetric) *touch*. These are verbs like *to be tied to*, *to stick to*, *to cling to*, *to clutch*, *to grip*, *to support* but also verbs that require an instrument such as *to be handcuffed to* and *to be fastened to*. This class of verbs can be tested using drawings, comparable to the ones that we used in the current experiment.

Such additional experiments that test more symmetric verb can show not only whether a linear situation is acceptable and whether it is preferred over an open circle, but also whether the open circle in and on itself is acceptable for a reciprocal sentence. For the verb *hold hands*, we found that the open circle, even though it was less preferred, was still accepted 44% of the time. This is quite a high number, especially compared to the other open circle test items. Testing more verbs can show us whether this is a finding that holds for all symmetric verbs or not. If it is indeed the case that an open circle is also acceptable even though consistently to a lesser extent than a linear situation, this would tell us that there might be more to the theory of reciprocals in general than we initially thought. We would then need to raise new questions and look into what it means for one situation to be a better reciprocal situation than the other – while both are being accepted.

There is also another direction for further research. Testing a large number of symmetric verbs can surely tell us more about geometry-sensitivity and whether our revised hypothesis holds for these verbs. On the other hand, we can also test the predictions that the hypothesis makes for non-symmetric verbs. As we mentioned, we suspect that the linear situations that we tested in our experiment were not maximal for the non-symmetric verbs, because subjects also considered alternatives with different directionality of actions. We can test whether this explanation holds by giving subjects an acceptability judgement task with those alternatives. If they do indeed accept them, then we strengthen the revised hypothesis

that says that we do consider alternatives with different directionality of actions, as was explained above.

On a final note, we may wonder what happens with verbs that do not presuppose any spatial configuration in the first place (e.g. *know*, *admire*). Obviously geometry-sensitivity is not applicable here, but a question that we may wish to ask is whether there is a more general mechanism that surfaces as geometry-sensitivity for action verbs, but as some other principle for other verbs. Research on different types of verbs is needed to give us more insight into this matter.

Chapter 6

Conclusion

In this thesis, a geometry-sensitive hypothesis for reciprocals was put forward. According to that hypothesis, the spatial configuration of the reciprocal antecedent set is a factor that needs to be taken into account when interpreting a reciprocal sentence. This would indicate that any account that deals with reciprocals needs to be sensitive to geometry. By proposing such a hypothesis, we aimed to give an explanation for the fact that some reciprocal sentences receive an interpretation that appears at first sight to be non-maximal. Because a geometry-sensitive hypothesis determines maximality given a geometrical configuration, it explains how such interpretations are possible. In order to test the hypothesis, we conducted an experiment in which we compared situations that differ merely in spatial configuration. According to a geometry-sensitive hypothesis, these situations are not equivalent, while any hypothesis that does not care about geometry would treat them as such.

The overall results did not support a general geometry-sensitive hypothesis. However, a few outliers in the results suggest that in some cases geometry might indeed be a relevant factor. These exceptions led the way towards a refined version of the initial hypothesis, and we have reason to believe that geometry-sensitivity in the revised formulation can be supported once we test an extended class of verbs. Additional research is needed to test this and should investigate also whether more factors like “geometry” exist. Once we have a clearer picture of all factors that influence the interpretation of a reciprocal sentence, we may get closer to a theory that captures the large variety of reciprocal interpretations.

References

- Dalrymple, M., Kanazawa, M., Kim, Y., Mchombo, S. & Peters, S. (1998). Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21, 159-210.
- Kański, Z. (1987). Logical Symmetry and Natural Language Reciprocals. In I. Ruzsa and A. Szabolcsi (eds.), *Proceedings of the 1987 Debrecen Symposium on Logic and Language* (49-68), Budapest, Akadémiai Kiadó.
- Kerem, N., Friedmann, N. & Winter, Y. (2009). Typicality Effects and the Logic of Reciprocity. In S. Ito and E. Cormany (eds.), *Proceedings of Semantics and Linguistic Theory, SALT XIX*, Cornell University.
- Langendoen, D. T. (1978). The Logic of Reciprocity. *Linguistic Inquiry* 9(2), 177-197.
- Laurence, S. & Margolis, E. (1999). Concepts and Cognitive Science. In E. Margolis and S. Laurence (eds.), *Concepts: Core Readings* (3-81). Cambridge, MA: MIT Press.
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T.E. Moore (eds.), *Cognitive development and the acquisition of language*, New York: Academic Press, 111-144.
- Rosch, E. & Mervis, C.B. (1975). Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology* 7(4), 573-605.
- Sabato, S. The Semantics of Reciprocal Expressions in Natural Language. Msc. Thesis.
- Sabato, S. & Winter, Y. (2011). Relational Concepts and the Logic of Reciprocity. Unpublished manuscript.
- Smith, E.E., Shoben, E.J. & Rips, L.J. (1974). Structure and Process in Semantic Memory: A Featural Model for Semantic Decisions. *Psychological Review* 81(3), 214-24.

Appendix A

A1: Order of questions in pilot study 3 questionnaires

Version 1	Version 2
Part 1: Yes/No	Part 1: Yes/No
1. Lean on (line) 2. Wipe (act on self) - filler 3. Comb (OC) 4. Paint (one relation) - filler 5. Point at (line) 6. Lean on (OC) 7. Scrape (closed circle) - filler 8. Point at (OC) 9. Push (one relation) - filler 10. Comb (line)	1. Pull (OC) 2. Wipe (act on self) – filler 3. Shake (line) 4. Paint (one relation) – filler 5. Stab (OC) 6. Pull (line) 7. Scrape (closed circle) – filler 8. Stab (line) 9. Push (one relation) – filler 10. Shake (OC)
Part 2: Forced choice	Part 2: Forced choice
11. Pull (OC vs. line) 12. Point at (open circle M vs. F) - filler 13. Shake (line vs. OC) 14. Scrape (line vs. closed circle) - filler 15. Stab (OC vs. line)	11. Lean on (line vs. OC) 12. Point at (open circle M vs. F) - filler 13. Comb (OC vs. line) 14. Scrape (line vs. closed circle) - filler 15. Point at (line vs. OC)

Leeftijd:

Geslacht: M / V *

Opleiding:

Jaar:

Moedertaal:

Links- of rechtshandig: L / R *

* Omcirkel het antwoord dat van toepassing is.

Deel 1 (van 2)

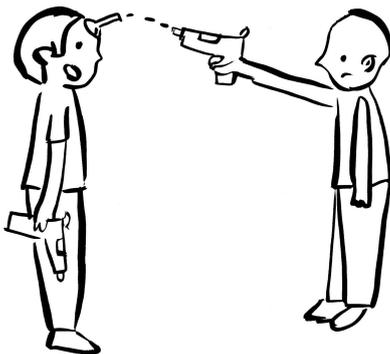
Let op: Lees deze instructie goed door voordat je begint met de vragenlijst

Het eerste deel van deze vragenlijst bestaat uit 10 korte vragen, waarbij we geïnteresseerd zijn in de verschillende betekenissen van werkwoorden. Elke vraag bestaat uit een tekening en een stelling, en we willen jouw **mening**:

Is de afgebeelde gebeurtenis in de tekening volgens jou wel of niét een mogelijke uitbeelding van de stelling?

Hieronder volgen twee simpele voorbeelden. In de tekeningen zie je twee jongens: de jongen rechts schiet op de jongen links, maar hoewel de jongen links ook een pistool heeft, schiet hij niet.

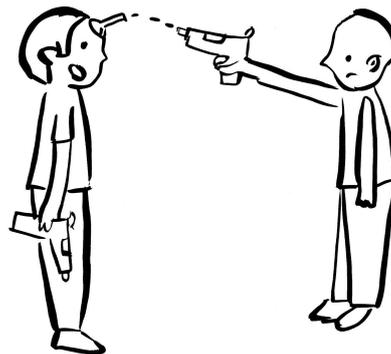
Voorbeeld 1:



In deze tekening schiet de jongen rechts op de jongen links

Ja / Nee

Voorbeeld 2:



In deze tekening schiet de jongen links op de jongen rechts

Ja / Nee

De stellingen in de vragenlijst zullen ingewikkelder zijn, en meestal is er geen duidelijk “goed” of “fout” antwoord. Geef in dat geval jouw eigen mening. Denk bij elke stelling niet te lang na, het gaat om je **eerste ingeving**. Je hebt ongeveer 3 minuten om de 10 vragen te beantwoorden.



1. In deze tekening leunen Anne, Joyce en Marlies op elkaar

Ja / Nee

(Omcirkel het antwoord van je keuze)



2. In deze tekening drogen Ilse, Lotte en Tessa elkaar af

Ja / Nee

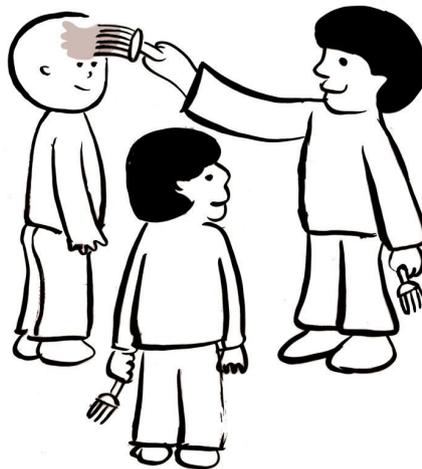
(Omcirkel het antwoord van je keuze)



3. In deze tekening kammen Cindy, Sofie en Emma elkaar

Ja / Nee

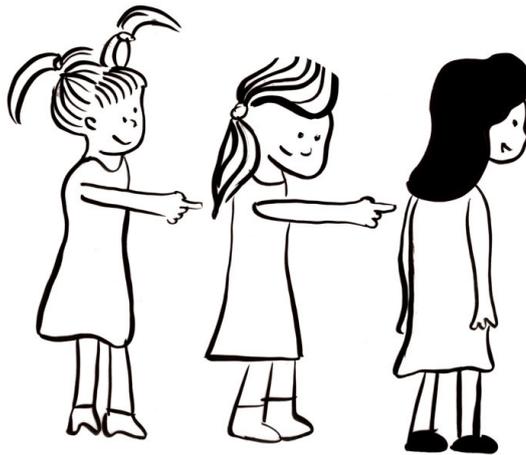
(Omcirkel het antwoord van je keuze)



4. In deze tekening verven Bas, Daan en Arnout elkaar

Ja / Nee

(Omcirkel het antwoord van je keuze)



5. In deze tekening wijzen Eva, Nynke en Laura naar elkaar

Ja / Nee

(Omcirkel het antwoord van je keuze)



6. In deze tekening leunen Anne, Joyce en Marlies op elkaar

Ja / Nee

(Omcirkel het antwoord van je keuze)



7. In deze tekening krabben Rachel, Marie en Irma elkaar

Ja / Nee

(Omcirkel het antwoord van je keuze)



8. In deze tekening wijzen Eva, Nynke en Laura naar elkaar

Ja / Nee

(Omcirkel het antwoord van je keuze)



9. In deze tekening duwen Bert, Felix en Daniel elkaar

Ja / Nee

(Omcirkel het antwoord van je keuze)



10. In deze tekening kammen Cindy, Sofie en Emma elkaar

Ja / Nee

(Omcirkel het antwoord van je keuze)

EINDE DEEL 1

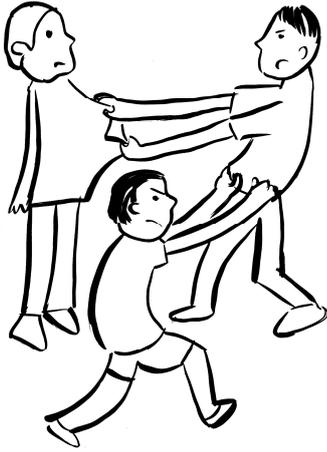
Deel 2 (van 2)

Let op: Lees deze instructie goed door voordat je verder gaat

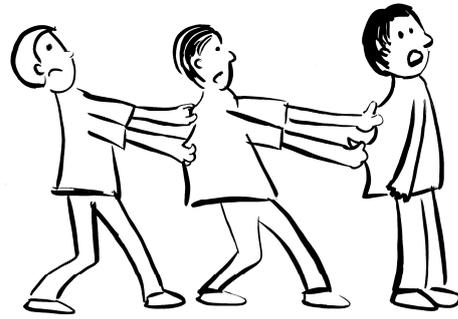
Het tweede deel van deze vragenlijst bestaat uit 5 korte vragen, waarbij we wederom geïnteresseerd zijn in de verschillende betekenissen van werkwoorden. Elke vraag bestaat nu uit twee tekeningen en een stelling, en we willen jouw **mening**:

Welke van de twee tekeningen is volgens jou de beste uitbeelding van de stelling?

Net als in deel 1 is er meestal geen duidelijk “goed” of “fout” antwoord. Ook al vind je het een moeilijke keuze, kies toch altijd één van de twee opties. Denk bij elke stelling niet te lang na, het gaat weer om je **eerste ingeving**. Je hebt ongeveer 2 minuten om de 5 vragen te beantwoorden.



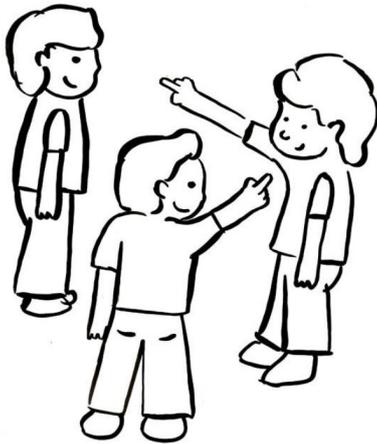
A



B

11. In deze tekening trekken Gijs, Lucas en Thomas aan elkaar

(Omcirkel de letter die hoort bij de beste optie)



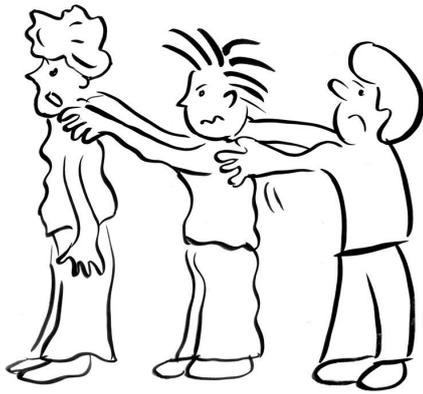
A



B

12. In deze tekening wijzen Tim, Kees en Edwin naar elkaar

(Omcirkel de letter die hoort bij de beste optie)



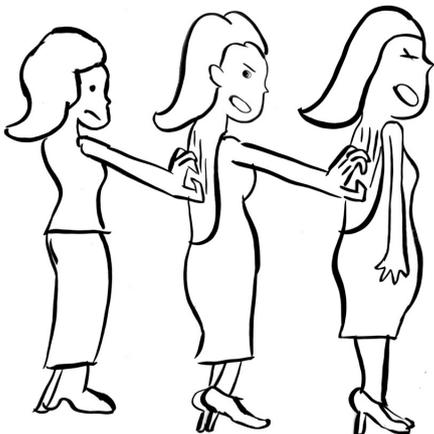
A



B

13. In deze tekening schudden Peter, Stan en Frank elkaar

(Omcirkel de letter die hoort bij de beste optie)



A



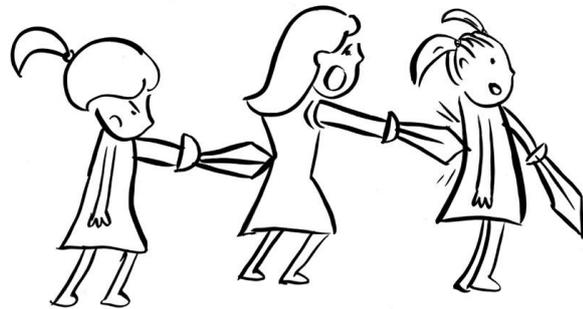
B

14. In deze tekening krabben Rachel, Marie en Irma elkaar

(Omcirkel de letter die hoort bij de beste optie)



A



B

15. In deze tekening steken Nadien, Katja en Suzan elkaar

(Omcirkel de letter die hoort bij de beste optie)

EINDE DEEL 2

Tot slot hebben we een paar vragen voor je:

Deel 1 (één plaatje)

1. Vond je de taak moeilijk?
2. Heb je tijdens het invullen een strategie gebruikt? Zo ja, welke?
3. Heb je enig idee waar deze taak over ging? Zo ja, waarover denk je?

Deel 2 (twee plaatjes)

1. Vond je de taak moeilijk?

2. Heb je tijdens het invullen een strategie gebruikt? Zo ja, welke?

3. Heb je enig idee waar deze taak over ging? Zo ja, waarover denk je?

Algemeen: Heb je nog andere opmerkingen?

Hartelijk dank voor je deelname!

Appendix B

B1: Questionnaire structure A and B

Questionnaire A: Acceptability Verbs 1-11

Target Questions

1	Stab	Line	F
2	Stab	Open circle	F
3	Scrape	Line	F
4	Scrape	Open circle	F
5	Apply makeup	Line	M
6	Apply makeup	Open circle	M
7	Wipe	Line	F
8	Wipe	Open circle	F
9	Clean	Line	M
10	Clean	Open circle	M
11	Comb	Line	F
12	Comb	Open circle	F
13	Shake	Line	M
14	Shake	Open circle	M
15	Pinch	Line	F
16	Pinch	Open circle	F
17	Paint	Line	M
18	Paint	Open circle	M
19	Point at	Line	F
20	Point at	Open circle	F
21	Hit	Line	M
22	Hit	Open circle	M

Fillers

1	Hit	Closed circle	M
2	Apply makeup	Closed circle	M
3	Stab	Closed circle	F
4	Comb	One relation	F
5	Paint	One relation	M
6	Pinch	Long line	F
7	Clean	Big open circle	M
8	Scrape	Big open circle	F
9	Wipe	Act on self	F
10	Point at	Open circle (give verb shake)	M
11	Shake	Open circle (give verb point)	F

Order

1A1

1A2

1	Comb	Line	OC	<i>F</i>
2	<i>Pinch</i>	<i>Long line</i>		<i>F</i>
3	Paint	OC	Line	<i>M</i>
4	<i>Point at</i>	<i>Wrong verb</i>		<i>M</i>
5	Stab	Line	OC	<i>F</i>
6	<i>Clean</i>	<i>Big open circle</i>		<i>M</i>
7	Hit	OC	Line	<i>M</i>
8	Scrape	Line	OC	<i>F</i>
9	Makeup	Line	OC	<i>M</i>
10	<i>Wipe</i>	<i>Acting on self</i>		<i>F</i>
11	Shake	OC	Line	<i>M</i>
12	Pinch	OC	Line	<i>F</i>
13	Point at	OC	Line	<i>F</i>
14	Clean	Line	OC	<i>M</i>
15	<i>Comb</i>	<i>One relation</i>		<i>F</i>
16	Stab	OC	Line	<i>F</i>
17	Makeup	OC	Line	<i>M</i>
18	Paint	Line	OC	<i>M</i>
19	<i>Shake</i>	<i>Wrong verb</i>		<i>F</i>
20	Hit	Line	OC	<i>M</i>
21	<i>Stab</i>	<i>Closed circle</i>		<i>F</i>
22	Pinch	Line	OC	<i>F</i>
23	Scrape	OC	Line	<i>F</i>
24	Wipe	Line	OC	<i>F</i>
25	<i>Makeup</i>	<i>Closed circle</i>		<i>M</i>
26	Shake	Line	OC	<i>M</i>
27	Comb	OC	Line	<i>F</i>
28	<i>Paint</i>	<i>One relation</i>		<i>M</i>
29	<i>Scrape</i>	<i>Big open circle</i>		<i>F</i>
30	Clean	OC	Line	<i>M</i>
31	<i>Hit</i>	<i>Closed circle</i>		<i>M</i>
32	Point at	Line	OC	<i>F</i>
33	Wipe	OC	Line	<i>F</i>

- All verbs occur 3 times
- Fillers are *italic*
- Two versions (mirrored target questions)

Questionnaire B: Acceptability Verbs 12-22

Target Questions

1	Push	Line	M
2	Push	Open circle	M
3	Lean on	Line	F
4	Lean on	Open circle	F
5	Pull	Line	M
6	Pull	Open circle	M
7	Touch	Line	F
8	Touch	Open circle	F
9	Tickle	Line	F
10	Tickle	Open circle	F
11	Pat on back	Line	F
12	Pat on back	Open circle	F
13	Kick	Line	M
14	Kick	Open circle	M
15	Hold hands	Line	F
16	Hold hands	Open circle	F
17	Shoot	Line	M
18	Shoot	Open circle	M
19	Play [sound]	Line	M
20	Play [sound]	Open circle	M
21	Wet	Line	M
22	Wet	Open circle	M

Fillers

1	Play [sound]	Closed circle	F
2	Kick	Closed circle	M
3	Pat on back	Closed circle	F
4	Push	One relation	M
5	Lean on	One relation	F
6	Pull	Long line	M
7	Touch	Long line	F
8	Hold hands	Big open circle	F
9	Tickle	Open circle (give verb shoot)	M
10	Shoot	Open circle (give verb wet)	F
11	Wet	Open circle (give verb tickle)	F

Order

		1B1	1B2		
1	Lean on	Line	OC		F
2	<i>Hold hands</i>	<i>Big open circle</i>			F
3	Play [sound]	OC	Line		M
4	<i>Wet</i>	<i>Wrong verb</i>			F
5	Pat on back	Line	OC		F
6	<i>Tickle</i>	<i>Wrong verb</i>			M
7	Kick	OC	Line		M
8	Pull	Line	OC		M
9	Push	Line	OC		M
10	<i>Touch</i>	<i>Long line</i>			F
11	Shoot	OC	Line		M
12	Hold hands	OC	Line		F
13	Wet	OC	Line		M
14	Tickle	Line	OC		F
15	<i>Lean on</i>	<i>One relation</i>			F
16	Pat on back	OC	Line		F
17	Push	OC	Line		M
18	Play [sound]	Line	OC		M
19	<i>Shoot</i>	<i>Wrong verb</i>			F
20	Kick	Line	OC		M
21	<i>Pat on back</i>	<i>Closed circle</i>			F
22	Hold hands	Line	OC		F
23	Pull	OC	Line		M
24	Touch	Line	OC		F
25	<i>Push</i>	<i>One relation</i>			M
26	Shoot	Line	OC		M
27	Lean on	OC	Line		F
28	<i>Play [sound]</i>	<i>Closed circle</i>			F
29	<i>Pull</i>	<i>Long line</i>			M
30	Tickle	OC	Line		F
31	<i>Kick</i>	<i>Closed circle</i>			M
32	Wet	Line	OC		M
33	Touch	OC	Line		F

- All verbs occur 3 times
- Fillers are *italic*
- Two versions (mirrored target questions)

Leeftijd:

Geslacht: M / V *

Opleiding:

Jaar:

Moedertaal:

Links- of rechtshandig: L / R *

* Omcirkel het antwoord dat van toepassing is.

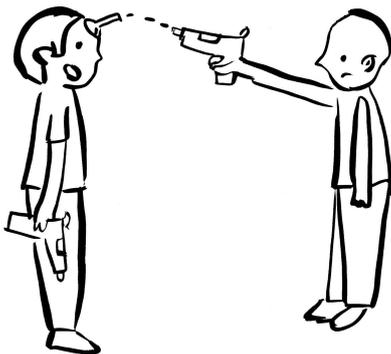
Let op: Lees deze instructie goed door voordat je begint met de vragenlijst

Deze vragenlijst bestaat uit 33 korte vragen, waarbij we geïnteresseerd zijn in de verschillende betekenissen van werkwoorden. Elke vraag bestaat uit een tekening en een stelling, en we willen jouw **mening**:

Is de afgebeelde gebeurtenis in de tekening volgens jou wel of niet een mogelijke uitbeelding van de stelling? Omcirkel het antwoord van je keuze.

Hieronder volgen twee simpele voorbeelden. In de tekeningen zie je twee jongens: de jongen rechts schiet op de jongen links, maar hoewel de jongen links ook een pistool heeft, schiet hij niet.

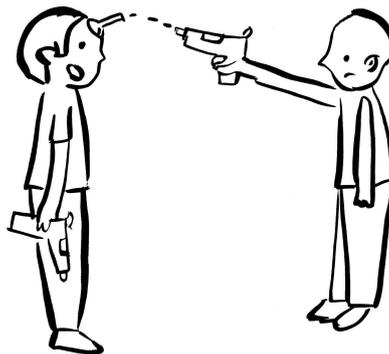
Voorbeeld 1:



In deze tekening schiet de jongen rechts op de jongen links

Ja / Nee

Voorbeeld 2:



In deze tekening schiet de jongen links op de jongen rechts

Ja / Nee

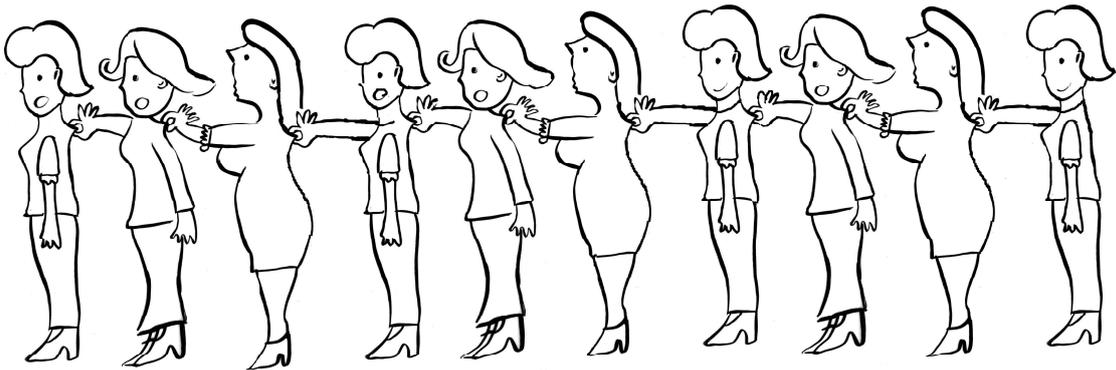
De stellingen in de vragenlijst zullen ingewikkelder zijn, en meestal is er geen duidelijk “goed” of “fout” antwoord. Geef in dat geval jouw eigen mening. Denk bij elke stelling niet te lang na, het gaat om je **eerste ingeving**. Je hebt ongeveer 10 minuten om de 33 vragen te beantwoorden. Daarna volgen er nog een paar afsluitende vragen.

Wacht a.u.b. met het omslaan van de pagina totdat de afnemer van de enquête het sein geeft!



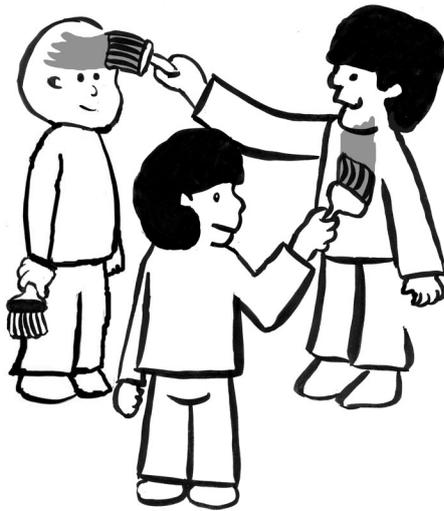
In deze tekening kammen Cindy, Sofie en Emma elkaar

Ja / Nee



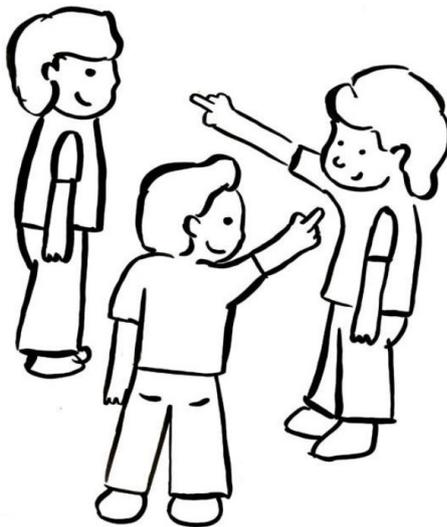
In deze tekening knippen de meisjes 1 t/m 10 elkaar

Ja / Nee



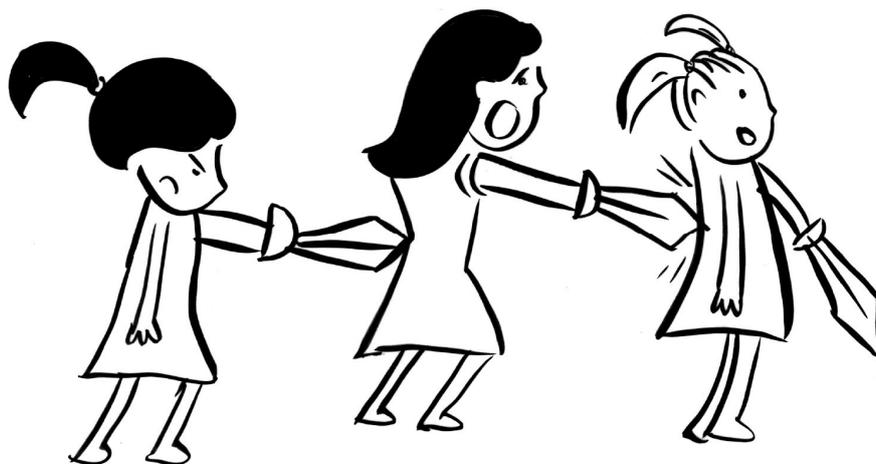
In deze tekening verven Bas, Daan en Arnout elkaar

Ja / Nee



In deze tekening schudden Tim, Kees en Edwin elkaar

Ja / Nee



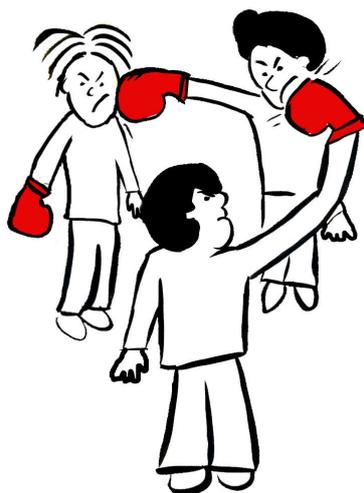
In deze tekening steken Nadien, Katja en Suzan elkaar

Ja / Nee



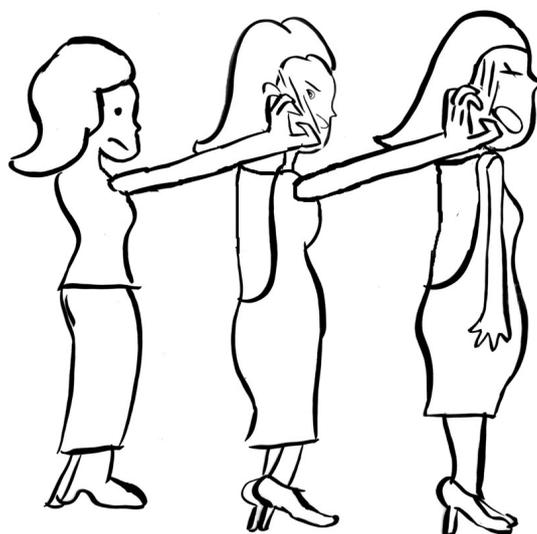
In deze tekening maken de jongens 1 t/m 10 elkaar schoon

Ja / Nee



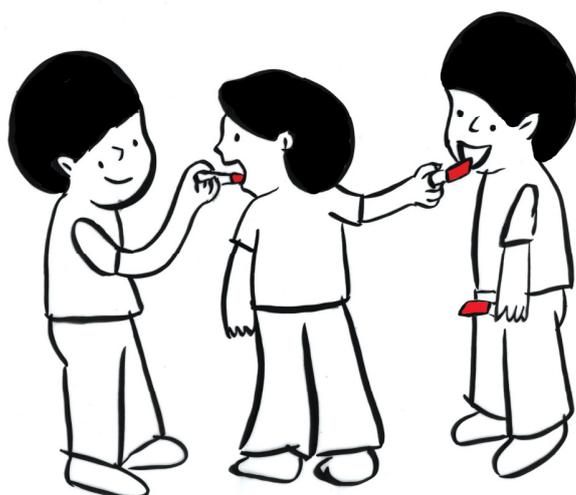
In deze tekening slaan Jaap, Tom en Jan elkaar

Ja / Nee



In deze tekening krabben Rachel, Marie en Irma elkaar

Ja / Nee



In deze tekening maken Erik, Dennis en Michel elkaar op

Ja / Nee



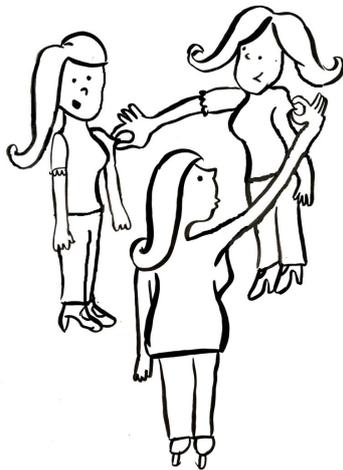
In deze tekening drogen Ilse, Lotte en Tessa elkaar af

Ja / Nee



In deze tekening schudden Peter, Stan en Frank elkaar

Ja / Nee



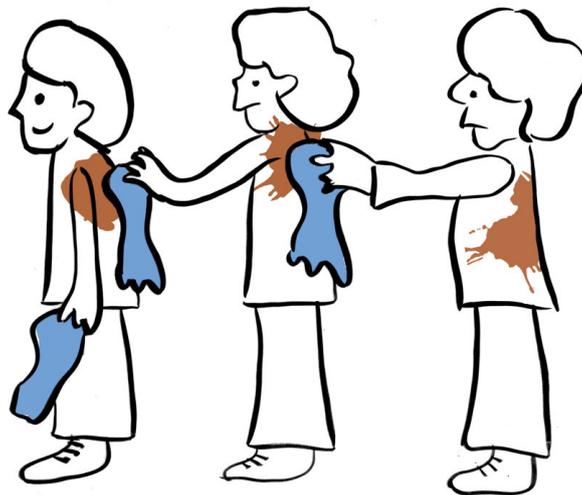
In deze tekening knijpen Daphne, Lisa en Sanne elkaar

Ja / Nee



In deze tekening wijzen Eva, Nynke en Laura naar elkaar

Ja / Nee



In deze tekening maken Mark, Rob en Olaf elkaar schoon

Ja / Nee



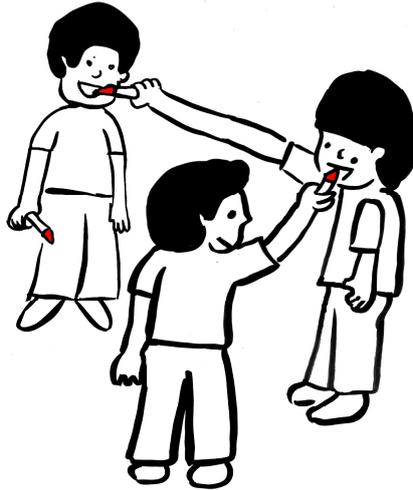
In deze tekening kammen Cindy, Sofie en Emma elkaar

Ja / Nee



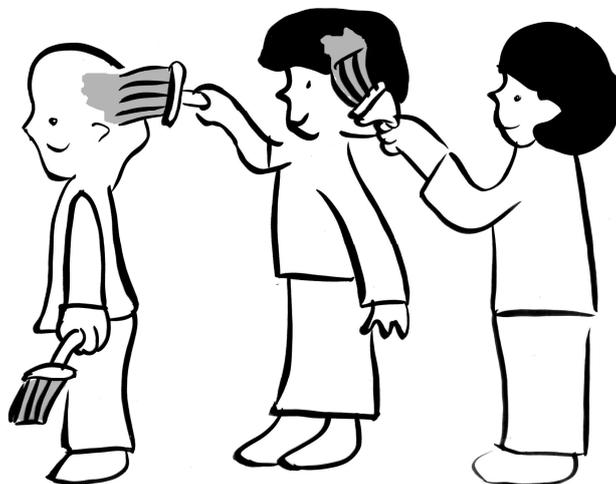
In deze tekening steken Nadien, Katja en Suzan elkaar

Ja / Nee



In deze tekening maken Erik, Dennis en Michel elkaar op

Ja / Nee



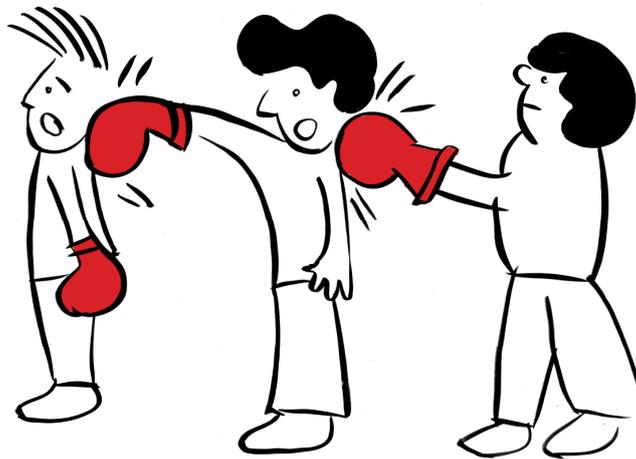
In deze tekening verven Bas, Daan en Arnout elkaar

Ja / Nee



In deze tekening wijzen Roos, Marja en Kim naar elkaar

Ja / Nee



In deze tekening slaan Jaap, Tom en Jan elkaar

Ja / Nee



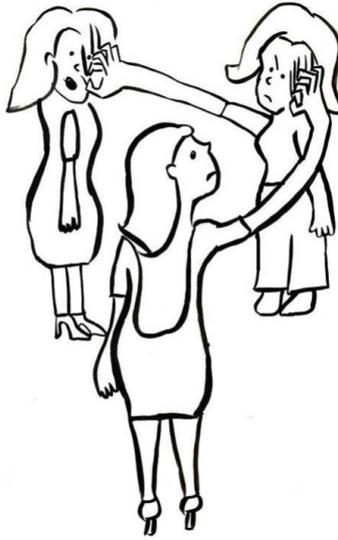
In deze tekening steken Nadien, Katja en Suzan elkaar

Ja / Nee



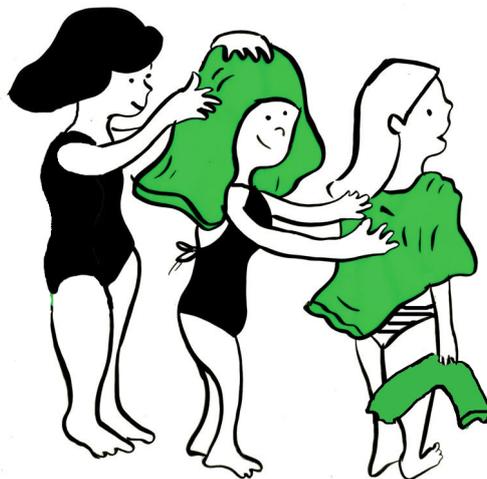
In deze tekening knijpen Daphne, Lisa en Sanne elkaar

Ja / Nee



In deze tekening krabben Rachel, Marie en Irma elkaar

Ja / Nee



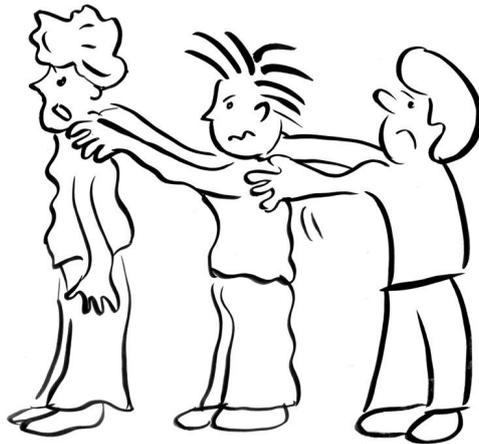
In deze tekening drogen Ilse, Lotte en Tessa elkaar af

Ja / Nee



In deze tekening maken Erik, Dennis en Michel elkaar op

Ja / Nee



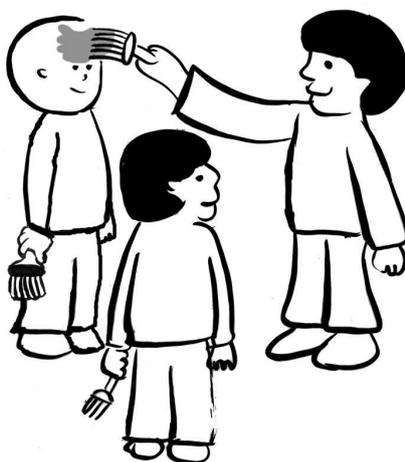
In deze tekening schudden Peter, Stan en Frank elkaar

Ja / Nee



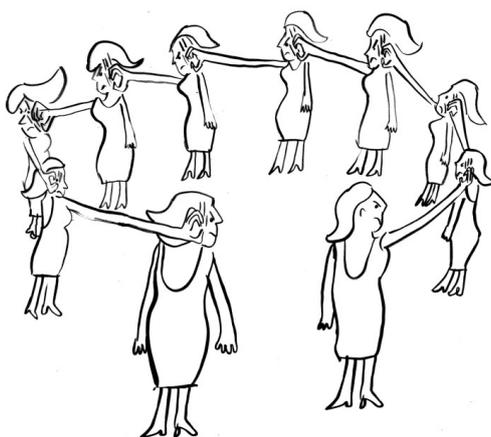
In deze tekening kammen Cindy, Sofie en Emma elkaar

Ja / Nee



In deze tekening verven Bas, Daan en Arnout elkaar

Ja / Nee



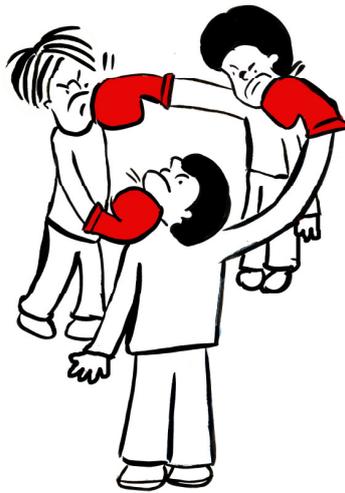
In deze tekening krabben de meisjes 1 t/m 10 elkaar

Ja / Nee



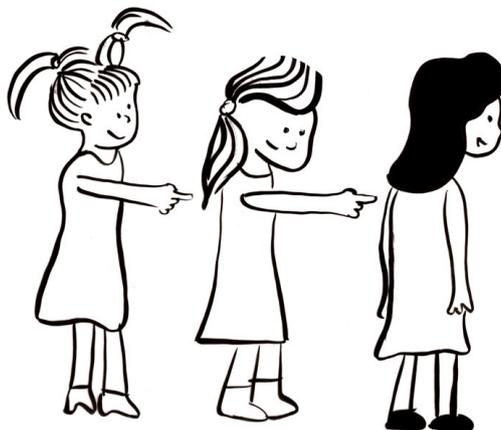
In deze tekening maken Mark, Rob en Olaf elkaar schoon

Ja / Nee



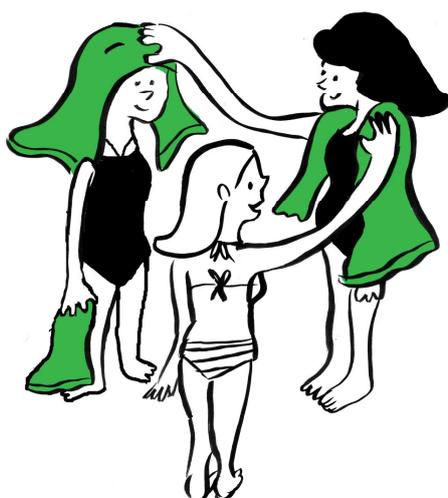
In deze tekening slaan Jaap, Tom en Jan elkaar

Ja / Nee



In deze tekening wijzen Eva, Nynke en Laura naar elkaar

Ja / Nee



In deze tekening drogen Ilse, Lotte en Tessa elkaar af

Ja / Nee

EINDE VRAGENLIJST

Op de volgende pagina staan een paar afsluitende vragen, neem hiervoor zoveel tijd als je nodig hebt.

4. Vond je de taak moeilijk?

5. Heb je tijdens het invullen een strategie gebruikt? Zo ja, welke?

6. Heb je enig idee waar deze taak over ging? Zo ja, waarover denk je?

7. Heb je nog andere opmerkingen?

Hartelijk dank voor je deelname!

