

RESEARCH

Reciprocal expressions and the Maximal Typicality Hypothesis

Eva B. Poortman¹, Marijn E. Struiksmā¹, Nir Kerem², Naama Friedmann³ and Yoad Winter¹

¹ Utrecht University, Trans 10, 3512 JK, Utrecht, NL

² Google Israel Ltd., Yigal Alon 98, Tel Aviv 6789141, IL

³ Tel Aviv University, Tel Aviv 69978, IL

Corresponding authors: Eva B. Poortman (e.b.poortman@uu.nl), Marijn E. Struiksmā (m.struiksmā@uu.nl)

In two experiments, we study the effects of verb concepts on the interpretation of reciprocal expressions in Dutch and Hebrew. One experiment studies Hebrew to test a previous account, the Strongest Meaning Hypothesis, which suggests that listeners resolve ambiguity in reciprocal sentences using the logically strongest meaning that is consistent with the context. The results challenge this proposal, as participants often adopt a weaker meaning than what the Strongest Meaning Hypothesis expects. We propose that these results reflect the sensitivity of reciprocal quantifiers to verb concepts, which is modelled by a new principle, the *Maximal Typicality Hypothesis* (MTH). For any given reciprocal sentence, the MTH specifies a *core situation*: the maximal situation that is also maximally typical for the verb concept. The MTH predicts reciprocal sentences to be maximally acceptable in the core situation and, under certain conditions, in situations that contain it, but substantially less acceptable in other situations. To test this prediction, we conducted a two-part experiment among Dutch speakers: (a) a membership test that ranks typicality preferences with different verbs; (b) a truth-value judgement test with reciprocal sentences containing these verbs. The results show that the typical number of patients per agent varies between verbs, with a significant effect of these preferences on reciprocal quantification: the stronger the verb concept's bias is for one-patient situations, the weaker is the interpretation of reciprocal sentences containing it. These results support the MTH as a basis for a general theory of reciprocal quantification.

Keywords: concepts; typicality effects; reciprocity; Strongest Meaning Hypothesis (SMH); Maximal Typicality Hypothesis (MTH)

1 Introduction

A well-known fact about reciprocal expressions like *each other* and *one another* is their sensitivity to linguistic and non-linguistic contextual parameters. Of these parameters, there is a special place for the meaning of lexical predicates with which reciprocals appear. When using a complex reciprocal verb phrase like *admire each other* or *chase each other*, the first linguistic factor that affects reciprocity is the verb's meaning. However, as with other content words, meanings of verbs like *admire* or *chase* are notoriously hard to define. One of the challenges for a semantic theory of reciprocity stems directly from this fuzziness. Despite the considerable efforts that have been invested in studying reciprocal expressions, all previous works analyze their quantificational effects as separate from the fuzzy aspects of predicate meaning. This separation between predicate meanings and quantificational processes has led to considerable empirical shortcomings of even the most precise theories of reciprocals (Dalrymple et al. 1998; Sabato and Winter 2012). As we will show, there

are intriguing semantic regularities in the way quantification with reciprocal expressions is affected by verb concepts. These regularities challenge previous approaches, and lead us to a new, experimentally informed, theory of reciprocals. The theory that we propose accounts for the facts that Dalrymple et al.'s proposal fails to explain, while preserving some of the empirical predictions and theoretical insights that it aims to articulate. Thereby, the proposed theory extends the analysis of quantificational reciprocity to a previously unstudied terrain between formal semantics and lexical semantics.

A major challenge for previous studies of reciprocals involves pairs of sentences like (1) and (2) below.

- (1) John, Bill and George know each other.
 (2) John, Bill and George are biting each other.

The logical structures of (1) and (2) are identical, consisting of three elements: a reciprocal expression (*each other*), an antecedent set of cardinality two or more (*John, Bill and George*), and a verb concept (*know/bite*). This structure provides speakers with a concise way of describing complex situations. For example, sentence (1) naturally describes a situation with six acquaintance relations:¹ John knows Bill, Bill knows John, John knows George, George knows John, Bill knows George and George knows Bill. We henceforth refer to any situation that demonstrates this full mutuality between three agents as an *S6 situation*. While the *S6* situation is salient for sentence (1), this is not the case for sentence (2): when hearing sentence (2), we first think of a situation with only three relations, where every man is biting only one other man. We refer to such situations as *S3 situations*.

Since Langendoen's (1978) early work, many studies of reciprocals in formal semantics have presented compelling evidence that reciprocal sentences such as (1) and (2) describe different situations. Some of these works concentrate on the way to derive meanings for reciprocal sentences with different syntactic structures (e.g. Heim, Lasnik and May 1991; Beck 2001). Other works focus on the use of contextual and lexical information for selecting between different meanings for reciprocal sentences as in (1) and (2) (Dalrymple et al. 1998; Sabato & Winter 2012; Mari 2013).² This paper develops the second line by experimentally studying the role of lexical information in the formal semantics of reciprocals.

In their influential work, Dalrymple et al. (1998) show a way to account systematically for the different interpretations of reciprocal sentences. Dalrymple et al. propose that each occurrence of a reciprocal expression denotes one of six different logical operators, encoding different strategies for categorizing situations. For instance, the operator of *Strong Reciprocity* (SR) categorizes situations where every member of the antecedent set is connected by the given relation to every other member. Thus, in sentences like (1) and (2) where the antecedent set consists of three members, applying SR results in selecting *S6* as

¹ Throughout this paper, we use the term "relation" informally, to refer to ordered pairs of individuals. Sets of such pairs (=standard *binary relations*) are informally referred to as "situations". This non-standard terminology is convenient for presentational purposes.

² Further work on reciprocals in Murray (2007; 2008) and Dotlačil (2013) discusses their behavior in discourse and suggests theories that substantially diverge from the standard proposals of Dalrymple et al. and Heim et al. that take reciprocals to be standard quantificational operators, possibly with an additional anaphoric element. Murray and Dotlačil's works do not systematically address the semantic/pragmatic factors that affect the choice of a specific reciprocal interpretation, and the issues they deal with are largely orthogonal to the main topic of the present work. The present paper develops the work in Kerem et al. (2009), which has also been followed in other areas: see Poortman's (2014; 2017a; b) experimental finding that the interpretation of plural predicate conjunctions as in *the shapes are small and big/blue and big* is governed by a principle that follows the Maximal Typicality Hypothesis proposed in the present paper. For other recent experimental and theoretical work which is relevant to Poortman's findings, see Poortman & Pylkkänen (2016), Lee (2017), Scontras & Goodman (2017) and Winter (2017).

the only possible situation where the sentence can be truthfully asserted.³ Other operators by Dalrymple et al. impose laxer logical requirements, which also admit other situations in addition to *S6*, especially the kind of situation we called *S3*.

To select between the six logical operators they propose, Dalrymple et al. introduce a principle that they call the *Strongest Meaning Hypothesis* (SMH). The selection relies on assumptions about the linguistic and non-linguistic context in which the reciprocal expression is used. For each occurrence of a reciprocal expression in a given context, the SMH selects the operator that results in the *strongest sentential meaning* that is consistent with that context. Other operators are ruled out. For example, in sentence (1), Dalrymple et al.'s account assumes that the context puts no restrictions on the number of possible acquaintances that any of the three persons may have.⁴ As a result, the SMH selects the strongest sentence meaning, where *each other* is interpreted as the SR operator: every man knows every other man. This analysis predicts that sentence (1) can only be true in *S6*, which is in line with many speakers' intuitions. Sentence (2) is also successfully analyzed by the SMH, by making the plausible assumption that each individual can only bite one other individual at a time. Taking this assumption on the meaning of *bite* to be part of the context of (2), the SMH cannot select SR for the reciprocal expression in (2), as this would be inconsistent with the contextual information. Consequently, a logically weaker operator than SR is selected (Dalrymple et al.'s operator of *Intermediate Reciprocity*). This operator correctly analyzes sentence (2) as true in *S3* situations such as the one where John bites Bill, Bill bites George and George bites John, i.e. where each individual only bites one other individual.

We see that based on plausible contextual assumptions, the SMH correctly accounts for the intuitive distinction between sentences (1) and (2). This account relies on the assumption that the verb *bite* restricts the number of possible patients per agent, whereas the verb *know* does not. More generally, in all the cases that Dalrymple et al. discuss, their analysis assumes that the context categorically restricts the number of patients that agents may possibly have simultaneously, or the number of agents that patients may be connected to. This assumption is needed in order for the SMH to allow reciprocal meanings that are weaker than Strong Reciprocity. In sentence (2), this assumption is innocuous enough: it looks reasonable to assume that ordinary contexts categorically rule out situations where some agent bites more than one patient simultaneously. However, with many other sentences, this kind of categorical judgement is questionable. For instance, let us consider sentence (3) below.

(3) John, Bill and George are pinching each other.

In terms of the situations that it supports, sentence (3) does not differ much from sentence (2): intuitively, both sentences seem equally true in *S3* situations. However, despite this similarity in the sentential interpretations of (2) and (3), the predicates *pinch* and *bite* are quite different in terms of the meaning restrictions that they induce on possible contexts. While in natural contexts, biting two people simultaneously seems impossible, pinching two patients simultaneously is physically possible. As a result, in normal contexts where people use both hands, the SMH would select SR as the meaning of the reciprocal in (3), hence the SMH expects the sentence to only be acceptable in the *S6* situation. This

³ For precision, we should note that *S6* actually represents a class of situations where each agent acts on both other agents, and possibly itself. Since this last formal detail is irrelevant for our discussion, we sloppily refer to *S6* as one situation.

⁴ For consistency with Dalrymple et al.'s terminology, we refer here to the kind of world knowledge about predicates like *know* and *bite* as being part of the "context". Whether this term is adequate depends on one's conception of lexical meaning and pragmatic context, but it is irrelevant for our purposes.

prediction of the SMH seems problematic: intuitively, we believe that sentence (3) should be judged acceptable in *S3*. As we will see, this intuition is supported by our experimental findings.

When addressing this challenge for the SMH, we might choose one of two opposing lines:

1. In an attempt to salvage the SMH, we might assume that for many speakers, the context does in fact put restrictions on the number of possible patients that may be pinched simultaneously, similarly to how the number of possible patients is restricted for *bite* in (2). In such contexts, the SMH would again select a weaker logical operator than SR. This would license sentence (3) in *S3* situations.
2. An opposite approach would be to conjecture that reciprocals allow weaker meanings than SR for all sentences. One candidate for such a meaning might be the one that Langendoen (1978) calls *Weak Reciprocity*. Without a selection principle like the SMH, the WR operator renders all reciprocal sentences like (1)–(3) acceptable in *S3* situations.

Approach 1 predicts that all speakers who accept sentences like (3) in *S3* situations reject them in *S6*. Approach 2 expects all speakers to also accept sentences like (1) in *S3* situations. As we will see, both predictions turn out to be problematic. To avoid these problems, we need to retain a contextually-informed principle for selecting reciprocal meanings, but the selection process itself must be based on more fine-grained, experimentally testable, assumptions about the lexical semantics of predicates.

Developing this idea, we propose a new selection principle that we call the *Maximal Typicality Hypothesis* (MTH).⁵ Avoiding sharp categorical judgements about concepts as in Dalrymple et al.’s proposal, the MTH predicts truth-values for reciprocal sentences in correlation to a fuzzy measure of conceptual semantics: the *typicality* of different situations for verb concepts like *know*, *pinch* or *bite*. For any given reciprocal sentence, the MTH designates a so-called *core situation*: a situation that contains enough relations between the agents in the sentence to satisfy reciprocity. What we define as “enough relations” depends on typicality information about the verb concept. In sentence (1), only the maximum of six acquaintance relations between the men would be considered enough, since the concept *know* imposes no typicality restrictions with respect to the number of simultaneous patients per agent. Thus, the core situation for (1) is the maximal one, where every man knows every other man (*S6*). This is the only situation in which sentence (1) is expected to be fully acceptable. By contrast, in (2) and (3), both verb concepts have clear typicality preferences with respect to the number of simultaneous patients per agent. Having six relations (i.e. two per agent) as in *S6* would be against the typicality preferences of the verb concept (in the case of *pinch*), or would not even be considered an instance of the verb concept (in the case of *bite*). In such cases, the core situation that the MTH selects is the maximal situation with no more than one patient per agent, hence *S3* is selected as core situation for (2) and (3). Accordingly, the MTH expects both (2) and (3) to be true and fully acceptable in *S3*.

This qualitative analysis is tested in two experiments, which are reported in Sections 3 and 4. Furthermore, the MTH makes novel and more precise quantitative predictions about the relations between reciprocity and typicality. Since acceptability of reciprocal sentences is predicted on the basis of typicality preferences with verb concepts,

⁵ The MTH was first proposed in Kerem et al. (2009). The present work develops this idea and tests it more extensively.

the MTH expects that the lower the typicality of situations with two patients per agent, the higher the prominence of readings that are weaker than SR. The experiment reported in Section 4 tests this predicted correlation.

The paper is structured as follows. Section 2 introduces the MTH in detail, as an alternative to the SMH. Sections 3 and 4 report two experiments that tested the aforementioned predictions of the MTH against those of the SMH. Section 5 concludes the paper.

2 The Maximal Typicality Hypothesis

This section introduces the Maximal Typicality Hypothesis as a solution to the problems that are raised by the different acceptability patterns for sentences like (1)–(3). After some background on typicality in subsection 2.1, the MTH is introduced in section 2.2. Further, subsections 2.3–2.5 explain our method for testing the MTH.

2.1 Concept typicality

In terms of theories of mental concepts (Margolis & Laurence 1999), Dalrymple et al.'s formulation of the SMH only takes into account “sharp” aspects of the meaning of verb concepts such as *know*, *bite*, and *pinch*: whether certain situations – e.g. those compatible with SR – are possible or impossible in a given context. As mentioned above, such sharp distinctions do not easily account for the intuitive contrasts between sentences like (1), (2), and (3). To overcome this shortcoming of the SMH, the proposed Maximal Typicality Hypothesis revises the SMH by on the basis of the typicality preferences of verb concepts.

Since the 1970's, a host of psychological studies has shown that subjects consistently rank some instances of a one-place predicate concept as more typical than others (e.g. Rosch 1973; Smith et al. 1974; Rosch and Mervis 1975). For example, besides being able to categorize sparrows and ostriches within the *bird* category, and koalas and crocodiles outside of it, subjects also distinguish between members of a category: e.g. when subjects are asked to rank bird instances, sparrows are judged as more typical for the concept *bird* than ostriches. These rankings correlate with other measures of typicality, such as categorization speed (more typical instances are categorized faster than less typical ones) and error rate (more typical instances lead to fewer categorization errors than less typical ones). Throughout this paper, we will use the term ‘typicality effect’ to refer to this basic behavioral phenomenon about categorization.

While many nouns categorize simple entities, verbs categorize more complex *situations*: events and states containing different entities as participants. As we will show, verb concepts exhibit typicality effects with situations, similarly to the typicality effects that noun concepts show with entities. For reciprocal sentences and verb concepts, taking typicality into account means changing perspectives about Dalrymple et al.'s notion of context. In addition to the definitional aspects that the SMH considers (can a given situation be categorized as an instance of a verb concept *X*?), our proposed account also has recourse to aspects of typicality (what preferences between situations does a verb concept *X* induce?). This allows us to take into account more factors that affect the interpretation of reciprocal sentences. As we saw, the sharp distinction that the SMH makes between possible and impossible situations forces it to choose SR as the meaning of the reciprocal expression in both sentences (1) and (3). This is intuitively questionable for (3). Under the theory we propose, the interpretation of the two sentences differs due to differences in typicality information between the concepts *know* and *pinch*. More generally, our theory uses experimental evidence on typicality to make predictions about the interpretations of reciprocal sentences. As we show, this leads to more fine-grained predictions, which make the correct distinctions between sentences like (1), (2) and (3).

2.2 The MTH: Connecting typicality with the interpretation of reciprocals

In the formal semantic literature on reciprocals, there are two general approaches that can be discerned. One approach, as in Langendoen (1978), Dalrymple et al. (1998), or Beck (2001), assumes systematic ambiguity of reciprocal expressions. Another approach, suggested by Roberts (1987), is that reciprocals are truth-conditionally vague, similarly to quantificational words like *many*, *most* or *enough*. The important contribution of Dalrymple et al.'s SMH is its ability to analyze ambiguity resolution systematically based on contextual information. Sabato and Winter (2012) propose to use Dalrymple et al.'s insights without assuming that reciprocals are ambiguous, but by letting them have a general functional meaning that takes lexical or contextual knowledge about predicate meanings as its argument. This semantic strategy is closer to Roberts' proposal in its aim to avoid ambiguity of reciprocals. However, like Dalrymple et al., Sabato and Winter also assume categorical distinctions between possible and impossible predicate denotations. What these previous approaches lack is a fine-grained notion of predicate meanings which takes typicality preferences into account, and uses it to analyze speaker judgements on sentences like (3). The present proposal follows Roberts's initial approach by incorporating typicality preferences into a theory that develops Sabato and Winter's version of the SMH.

The proposed Maximal Typicality Hypothesis (MTH) uses information about typicality of different situations with respect to a given verb concept P . On the basis of this information on P , the MTH singles out one situation as the core situation for reciprocal sentences containing P . This core situation is the basis for speakers' interpretation of reciprocal sentences.⁶ Above we intuitively described the core situation as the situation that contains *enough* relations between the agents to fully satisfy reciprocity. In terms of reciprocity, for sentence (1) with the verb *know*, situation S_6 has enough relations, whereas for sentence (3) with the verb *pinch*, already S_3 has enough relations. To be more precise, we need to define what we mean by *enough for reciprocity*, and how the two sentences differ in this respect. According to the MTH, the difference between (1) and (3) follows from the observed typicality difference between the verbs. With the verb *pinch* we cannot add relations to S_3 without reaching a situation where one agent pinches two patients. Such a situation would be atypical for the verb *pinch*. Because we cannot add relations to S_3 without a reduction in its typicality for *pinch*, we consider S_3 to have enough relations for sentence (3). By contrast, with the verb *know*, we can add relations to S_3 without any change in typicality for the verb concept. Thus, S_3 does not attain enough relations for sentence (1). In technical terms, we describe this difference by observing that S_3 is the *maximal situation* among the most typical situations for the verb *pinch*. By contrast, S_3 is not maximal among the most typical situations for the verb *know*. Accordingly, the MTH selects S_3 as the core situation for sentence (3), but not for sentence (1). In general, this process of selecting the core situation is defined below.

Maximal Typicality Hypothesis (MTH): For a reciprocal sentence with a verb concept P in the scope of the reciprocal expression, situation S_c is the *core situation* for the sentence iff S_c is maximal among the situations that are most typical for P .⁷

⁶ Our use of the term "interpretation" is meant to highlight our conviction that the semantic truth-conditional analysis of plural sentences, including reciprocal sentences, should be relativized to speakers' beliefs about the predicate concepts in such sentences, as well as other contextual parameters.

⁷ Note that in this analysis, there may be two or more situations that attain such a maximum, but without any of these situations containing the other. In such cases, the MTH would define more than one situation as a core situation. Such cases do not surface in our experiments. For this reason, here and henceforth we assume uniqueness of the core situation, and refer to it as 'the core situation'.

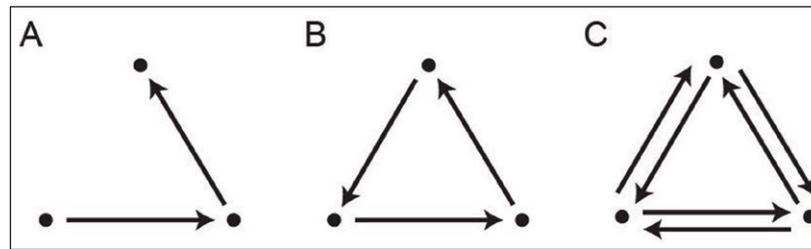


Figure 1: Three situations with three individuals.

The MTH assumes that similar to noun concepts, verb concepts invoke typicality judgements on situations.⁸ Formally (Winter 2017): the MTH defines the core situation as any maximal situation S_c among the situations S that attain a local maximum of the function $TYP_p(S)$ – the typicality of S for the predicate P . This typicality is a variable whose values can be experimentally estimated. Once we know which situations are most typical for a given verb concept, the MTH predicts the core situation for a reciprocal sentence containing that verb. The key for selecting the core situation among the most typical situations is the notion of *maximal situation*. This notion relies on our ability to order situations according to containment relations between them. For example, a situation like S_6 , where every agent acts on each of the two other agents, properly contains any S_3 situation. To illustrate this notion, it is helpful to consider the diagrams in Figure 1.

The arrows in Figure 1 represent directed actions like *pinch* between agents. Thus, Figures 1B and 1C represent S_3 and S_6 respectively. Figure 1A represents another kind of situation, to which we refer as S_2 . Since situation S_6 contains S_3 , and S_3 contains S_2 , we conclude that S_6 is maximal among the three situations in Figure 1. In sentence (1), we assume that S_6 , S_3 and S_2 are of equal typicality for the verb *know* (at least as far as the number of patients is concerned, see section 2.3 below). Among these three situations S_6 is the maximal, hence it is the one that is selected by the MTH. By contrast, in sentences (2) and (3), we assume that S_2 and S_3 are more typical for the verbs *bite* and *pinch* than S_6 (furthermore, S_6 may even be impossible with *bite*). The maximal situation among the most typical situations is now S_3 , and accordingly, this is the situation that is selected by the MTH as the core situation.

This selection of the core situation for reciprocal sentences like (1), (2) or (3) is the basis for explaining the acceptability pattern we observe with these sentences in all situations, including those that are not the core situation. Reciprocal sentences are always expected to be highly acceptable in the core situation. In addition to this core situation, there are two kinds of situations that we should consider:⁹

⁸ The verbs like *pinch* and *know* categorize events and states, respectively. The event/state distinction affects typicality preferences with verbs, hence the predictions of the MTH. However, since it does not directly affect the formulation of the MTH, we will not focus on it here. A similar point holds for tense: event verbs like *pinch* can take the progressive tense, while state verbs like *know* cannot. The choice of tense naturally affects typicality preferences. However, as we shall see, the MTH is about the correlation between typicality preferences and reciprocal interpretation. For this reason, holding the tense variable constant in the two measures does not affect the predicted correlation between them.

⁹ We here ignore situations that neither contain nor are contained by a core situation. As remarked in footnote 7, what we call *the* core situation should more accurately be called the *class* of maximal situations that attain maximal typicality. By definition, every other situation either contains or is contained by a situation in this class. For instance, S_3 situations were described as the situations where each of three participants only acts on one other participant. Assuming that participants do not act on themselves, there are six situations like that. Now, each of the situations where participants do not act on themselves either contains or is contained by one of these six core situations.

- a) *Situations that are properly contained in the core situation*: In those situations, reciprocal sentences are expected to be less acceptable than in the core situation, with decreasing acceptability the fewer relations there are. For example, the MTH defines *S6* as the core situation for sentence (1). Accordingly, sentence (1) is predicted to be less acceptable in *S3* than in *S6*, and less acceptable in *S2* than in *S3*.¹⁰
- b) *Situations that properly contain the core situation*: In those situations, there are “more than enough” relations to support reciprocity. Whether such situations are relevant for actual use of a given reciprocal sentence is determined by the felicity that speakers assign to them. For example, both sentences (2) and (3) are predicted to be highly acceptable in their core situation, *S3*. As for *S6* situations, judgements are affected by how felicitous situation *S6* is judged to be as a possible instance of the verb. Situation *S6* is usually judged to be a possible instance of the verb *pinch*. Accordingly, speakers are expected to judge sentence (3) as highly acceptable in *S6*. By contrast, *S6* situations are unlikely to be judged as acceptable for the concept *bite* to begin with. Accordingly, such situations are harder to use with sentence (2), and are often judged as unacceptable for this sentence.¹¹

2.3 Approximating typicality of complex situations

In order to test the MTH and its use with different situations, we first need information regarding the typicality of situations like *S2*, *S3* and *S6* for different verb concepts. The reason we consider situations like *S6* as atypical for the verb *pinch* is because agents in it pinch more than one patient simultaneously. Using this kind of typicality information, the predictions of the MTH are tested against acceptability judgements on reciprocal sentences in situations like *S2*, *S3* and *S6*. But how can we systematically gather experimental information about the typicality of situations like *S2*, *S3* and *S6* for different verb concepts? As was discussed in subsection 2.1, previous studies have measured typicality effects using standard tasks such as categorizing instances or ranking them with respect to how typical they are. Those works have mostly measured typicality effects for concepts expressed by nouns. For situations like *S2*, *S3* and *S6*, we approximate typicality by looking at their sub-situations in relation to typicality preferences of the verb. Consider for example the two situations in Figure 2, which depict *S3* and *S6* situations with the verb *pinch*.

Both situations show three women in pinching activities. In situation A, every woman is pinching only one other woman (*S3*), while in situation B, every woman is pinching every other woman (*S6*). Because of the agents using both hands in situation B, this situation is expected to be less typical for the verb *pinch* than situation A. When testing

¹⁰ Note that the MTH does not make any prediction about the rate in which acceptability declines. Thus, in principle, *S2* situations may be as (un)acceptable for (2)/(3) as they are for (1), despite the different core situations of these sentences (*S3* vs. *S6*, respectively). The question of how strongly acceptability declines in situations that are contained in the core situation is left for further research.

¹¹ An anonymous *Glossa* reviewer points out that this consideration about felicity of *S6* situations may give the impression that the appeal to typicality in our MTH proposal is dispensable. However, such an impression would be misleading. The MTH is a hypothesis about truth-conditions, based on typicality information. On top of any such theory of truth, we should always consider the plausibility of models that support sentences as true. For instance, the non-reciprocal sentences “Mary is pinching (biting) Dan and Max simultaneously” are expected to be true if Mary is pinching (biting) Dan while pinching (biting) Max at the same time. Due to the marginal status of the “simultaneous bite” scenario, the corresponding sentence is unacceptable. This kind of consideration is fairly standard, and is orthogonal to both the semantics of reciprocity and the phenomenon of typicality preferences.

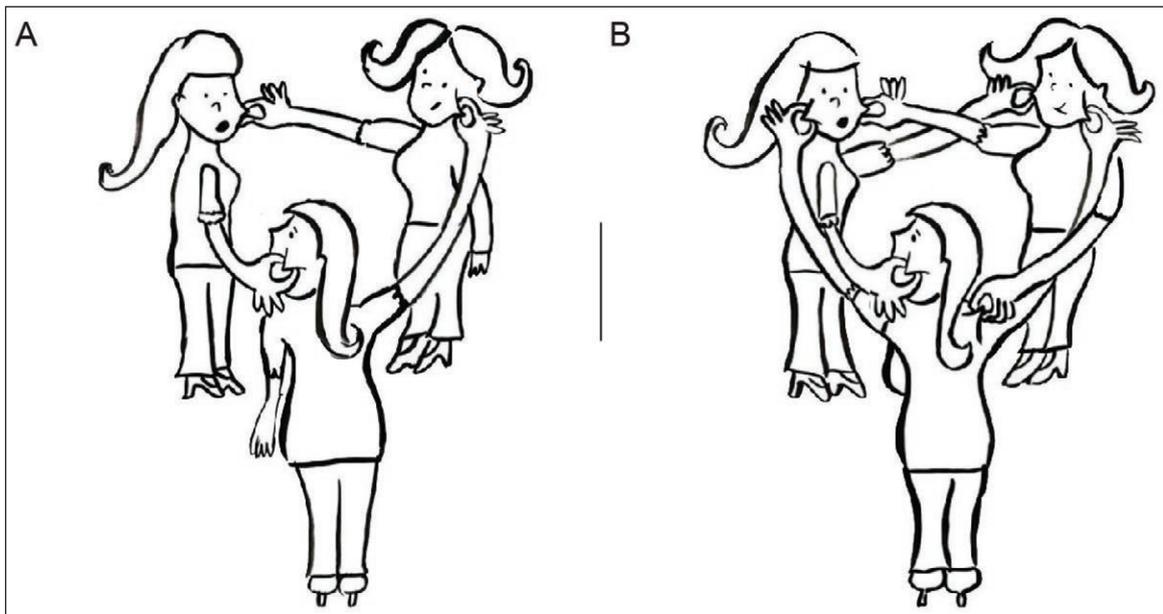


Figure 2: Two possible situations for the verb *pinch*: S3 (A) and S6 (B).

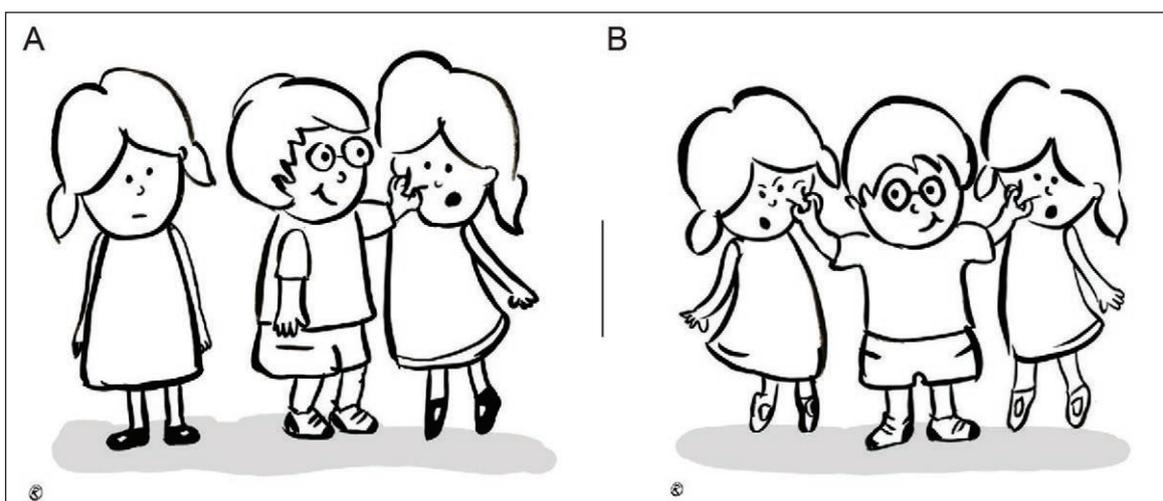


Figure 3: Two instances of the verb *pinch*: a one-patient situation (A) and a two-patient situation (B).

this expectation, we refer to the number of patients acted upon by a given agent as *patient cardinality*. For instance, in Figure 2, we describe the difference between the two situations by saying that the patient cardinality value is 1 (= one patient per agent) in Figure 2A, but 2 (= two patients per agent) in Figure 2B. This reduces the comparison between the two situations in Figure 2 to a comparison between the two situations in Figure 3. Our typicality experiments test which of the situations in Figure 3 is preferred as a more typical instance of the verb *pinch*. Based on the preference showed for Figure 3A over Figure 3B, we infer that S3 situations (Figure 1B) are more typical for the verb *pinch* than S6 situations (Figure 1C). This information is used for evaluating the MTH.

2.4 Predictions of the MTH

Based on the measurement of typicality effects described above in subsection 2.3, let us illustrate the precise empirical predictions of the MTH for sentences (1)–(3). For the verb *know* in (1), our experiments show that there is no typicality preference between simple

situations with different patient cardinality. This means that there is no difference in the representativeness of different instances of *knowing* for the concept, at least in terms of how many patients each agent knows. For instance, a state in which a person knows one other person is just as typical an instance of *know* as a state in which a person knows two other people. Using this measurement, we extrapolate that there is also no difference in typicality between reciprocal situations that differ merely in terms of patient cardinality (at least in terms of one patient vs. two). This means that the reciprocal situations *S2*, *S3* and *S6* are equally typical for the verb concept *know*. With this extrapolation about typicality of candidate situations, the MTH predicts that the core situation that (1) describes is *S6*: the maximal situation among the three situations *S2*, *S3* and *S6*. We thus predict that (1) is fully acceptable in *S6*. This agrees with the prediction of the SMH that the reciprocal in (1) is interpreted as the SR operator. Furthermore, we predict (1) to be less acceptable in *S3*, and even less so in *S2*, because these situations are properly contained in the core situation *S6*. This is a more fine-grained prediction than the truth/falsity prediction of the SMH, which expects (1) to be equally judged as false in both *S2* and *S3*.

In contrast to the verb *know*, the verb *bite* in (2) does show a typicality effect with respect to patient cardinality. In particular, instances of the concept *bite* that have one patient per agent are judged as being more typical than instances that have two patients per agent simultaneously.¹² We extrapolate from this that in the most typical reciprocal situations for the verb *bite*, there is no more than one patient per agent. The MTH predicts that the core situation described by sentence (2) is the maximal one among those situations. This is the situation in which each agent bites exactly one patient: *S3*. Adding more *biting* relations to such a situation would result in a situation that is not among the most typical situations. Thus, we predict (2) to be fully acceptable in *S3*, in line with the predictions of the SMH. Furthermore, we predict (2) to be unacceptable in *S6*: although the sentence is formally expected to be true in such a hypothetical situation, this scenario is unlikely to be a possible instance of the concept *bite* (see note 12). Finally, sentence (2) is expected to be less acceptable in *S2* than in the core situation *S3*, because *S2* is properly contained in this core situation.

The example that most clearly distinguishes the predictions of the MTH and the SMH is sentence (3). In this case, the core situation that is specified by the MTH does not coincide with the strongest possible interpretation that is selected by the SMH. As we saw, the SR meaning, where every boy is pinching every other boy, is physically possible for sentence (3). Therefore, according to the SMH, SR is the only possible reading for the reciprocal expression, and *S6* is expected to be the only situation in which sentence (3) is true. By contrast, the MTH predicts sentence (3) to be widely accepted in situation *S3*. As we will experimentally show, the concept *pinch* shows the expected typicality effect with respect to patient cardinality: a situation in which one agent pinches one patient is a more typical instance of *pinch* than a situation in which one agent pinches two patients simultaneously. In this sense, the verb *pinch* is similar to the verb *bite*, and we again extrapolate that the most typical situations for the verb *pinch* are those in which there is at most one patient per agent. From these situations, the MTH selects the maximal one, *S3*, as the core situation that the sentence describes. Accordingly, we expect sentence (3) to be fully acceptable in *S3*. In addition, since situation *S6* properly contains the core situation *S3*, we expect the sentence to be true in *S6*. Unlike *bite* in (2), the concept *pinch* allows *S6* as a possible, though atypical, instance of the concept, and hence we predict (3) to be acceptable in *S6*. The expectation that reciprocal sentences like (3) are judged

¹² In fact, the latter situation may even be considered physically impossible. This means that situations with multiple patients per agent may not be instances of the verb concept *bite*.

as acceptable in both *S3* and *S6* clearly distinguishes the predictions of the MTH from those of the SMH.

In order to test these predictions, we tested whether and how typicality preferences in terms of patient cardinality predict reciprocal interpretations using the MTH. As we will show, judgements on sentences like (3) reveal a systematic advantage of the MTH over the SMH.

2.5 Overview of experimental investigation

The experimental work presented in this paper had two goals. Firstly, we aimed to test interpretations of sentences like *the girls are pinching each other*, where the predictions of the SMH appear to be too strong. Secondly, we aimed to test whether the MTH makes the right predictions, with a focus on correlations between typicality preferences and reciprocal interpretation. The two experiments we report were designed to achieve these aims.

Experiment 1 tested the predictions of the SMH among Hebrew speakers. The experiment focused on action verbs in Hebrew like *cavat* ('pinch') that are expected to show a typicality preference for one-patient situations over two-patient situations. The materials of Experiment 1 were constructed based on a pretest that measured patient cardinality preferences for a large group of action verbs. The aim of the pretest was to select verbs for the main experiment that most clearly prefer one-patient situations over two-patient situations. The main experiment itself tested which situations are preferred for reciprocal sentences containing those action verbs that showed a preference for one-patient situations. Such a forced-choice task was used because challenging the SMH requires making sure that more than one situation is considered possible in the context of the sentence. For each sentence, participants chose between two realistically depicted situations with three human agents each: situation *S6*, in which every individual acts on every other individual (e.g. Figure 2B), and situation *S3* – one where every individual acts on exactly one other individual and is acted on by exactly one other individual (e.g. Figure 2A). The SMH predicts all sentences to be true in *S6* and false in *S3*, hence for *S6* to always be preferred over *S3*.

Next, Experiment 2 systematically tested the predictions of the Maximal Typicality Hypothesis: whether the core situation for a reciprocal sentence is maximal among those situations that are most typical for the verb concept in the reciprocal's scope. In order to have a broad variety of test cases for the MTH, Experiment 2 (unlike Experiment 1) included verbs that show different patient cardinality preferences. For logistic reasons, Experiment 2 was conducted with Dutch speakers. However, a pretest of Experiment 2 showed a significant correlation between Dutch and Hebrew in the typicality preferences of the verbs that were tested in Experiment 1. Thus, both experiments tested similar typicality effects in Dutch and Hebrew. In order to test the relation between reciprocal interpretation and typicality, Experiment 2 made use of two parts. Part 1 tested typicality effects for different types of verb concepts in isolation, specifically the preference for one-patient or two-patient situations. Part 2 of the experiment tested reciprocal interpretations, specifically the acceptability of reciprocal sentences in situations *S6* and *S3*. The MTH is evaluated on the basis of a correlation analysis between the two parts. Based on our measurement of typicality in part 1, we expected *S3* situations to be specified as the core situation in part 2 for sentences like *John, Bill and George are pinching each other*. A correlation analysis tested this expected relationship between the preference for a one-patient situation in part 1, and the acceptability of reciprocal sentences in the *S3* situation in part 2. In this way, Experiment 2 tested our MTH-based account and compared it to the predictions of the SMH.

3 Experiment 1: Testing the SMH

Experiment 1 studied Hebrew reciprocal sentences, testing the predictions of the SMH with verbs like *pinch* that are expected to show a preference for one-patient situations. A pretest collected examples of such verbs.

3.1 Pretest: Verbs with a one-patient preference

For a set of action verbs, the pretest measured whether there was a preference between situations with one patient per agent vs. situations with two patients per agent.

3.1.1 Participants

Fifty-three students from Tel Aviv University and Technion (Israel Institute of Technology) (14 female, age $M = 24$) participated as part of a class. All participants were native Hebrew speakers.

3.1.2 Materials and procedure

Thirty-two different Hebrew verbs were tested regarding their patient cardinality preferences in a pen-and-paper questionnaire. For twenty-five verbs, we expected a preference for a one-patient situation over a two-patient situation. For seven verbs, we expected a preference for the two-patient situation (see Table A1 in Appendix A). Each test item contained two instances of a verb, illustrated by two drawings showing a one-patient situation and a two-patient situation. Apart from the number of patients, the two situations were as similar as possible. An example is given in Figure 3, depicting two instances of the verb *covet*, ‘pinch’. All test items contained a verbal description of the situation in Hebrew: an agent-verb sentence (without a theme) referring to an activity associated with the verb concept, for example *ha-yeled covet*, ‘The boy is pinching’. Participants were instructed to indicate which of the two depicted situations “better describes the sentence”.¹³ In all items, due to different gender or age of the agent and patient(s), the subject of the sentence (a boy, girl, man, or woman) visibly referred to the agent in both drawings.

In addition, we included six filler items, which also contained two drawings. Their aim was to avoid automatic answers. The drawings in the filler items differed from one another with respect to a parameter other than the number of patients. For instance, one filler item showed a boy taking a photo in one drawing and merely holding a camera in the other drawing, accompanied by the Hebrew correlate of the sentence “The boy is taking a photo”. The task in the filler items was identical to the test items.

There were two versions of the questionnaire, with reversed order of items. In addition, in case the same verb was used for a test item and a filler item, the test item always preceded the filler item in both versions.

3.1.3 Results

The results of the pretest are given in Table A1 of Appendix A. Different verbs showed different preferences regarding patient cardinality. The proportion of participants who preferred the one-patient drawing over the two-patient drawing ranged from 8% to 90% between verbs. In general, most of the verbs showed a preference for one patient over two, indicating that participants considered the situation in which only one patient was

¹³ This phrasing is not perfect from a theoretical semantic point of view, but it proved clearest for the participants.

involved as a more typical instance of that verb than the situation with two patients. This is as expected, since we included mostly verbs that we believed would show a one-patient preference.

We used the results of the pretest to select verbs for the main part of the experiment, testing the SMH. The selection procedure is explained in the Materials section of section 3.2.

3.2 Experiment 1 (main part): Interpretation of reciprocal sentences

The main part of Experiment 1 measured the preferred interpretation of Hebrew reciprocal sentences containing verbs that showed a clear preference for one-patient situations in the pretest. We used a forced-choice task to measure which situation is preferred for the given reciprocal sentences: *S6* (where each individual is acting on every other individual) or *S3* (where each individual is acting on exactly one other individual and is acted on by exactly one other individual). In such contexts *S6* is visibly possible, hence according to the SMH, the sentences are expected to be uniformly judged as true in situation *S6* and false in *S3*. Thus, the SMH expects *S6* to be uniformly preferred over *S3*. The experiment tested this null hypothesis.

3.2.1 Participants

Fifty students from Tel Aviv University and Technion (Israel Institute of Technology) (20 female, age $M = 25$) participated, either as part of a class or for monetary compensation. All participants were native Hebrew speakers.

3.2.2 Materials and procedure

Eleven test item sentences were created based on the results of the pretest. We calculated the 95% confidence interval based on the preference for a one-patient situation from the pretest (C.I. 0.526–0.668). Using the upper boundary of the confidence interval analysis, we selected eleven verbs that showed highest preference for a one-patient situation and easily allowed graphical representation of reciprocal situations. We also included two control verbs that showed high preference for two-patient situations and allowed the required graphical representation. We then included the total of 13 verbs in reciprocal sentences of the form *A, B and C are P each other* (where *A, B* and *C* are proper names and *P* is a verb). An example for such a Hebrew sentence is in (4).

- (4) Dani, yoav ve-ro'i covtim ze-et-ze.
 Danny, Yoav and-Roy pinch-PRES.SG.MASC. this.ACC-this
 'Danny, Yoav and Roy pinch/are pinching each other.'

The resulting 13 sentences were tested for their interpretation in a forced-choice task with two drawings depicting different situations. For each of the sentences, there were two forced-choice trials. The main trial contained a drawing of *S6* (in which every individual acts on two other individuals) vs. a drawing of *S3* (in which every individual acts only on one other individual). The second trial served as a control, and contained again a drawing of *S3* vs. a drawing of situation *S2* (in which only two of the three individuals act on another individual). Apart from the number of actions, the two drawings were always as similar as possible, and the subject of the sentence clearly referred to the three agents in the drawings. The example trials for sentence (4) are given in Figure 4. Participants were instructed to indicate which of the two depicted situations better describes the sentence (see footnote 13).

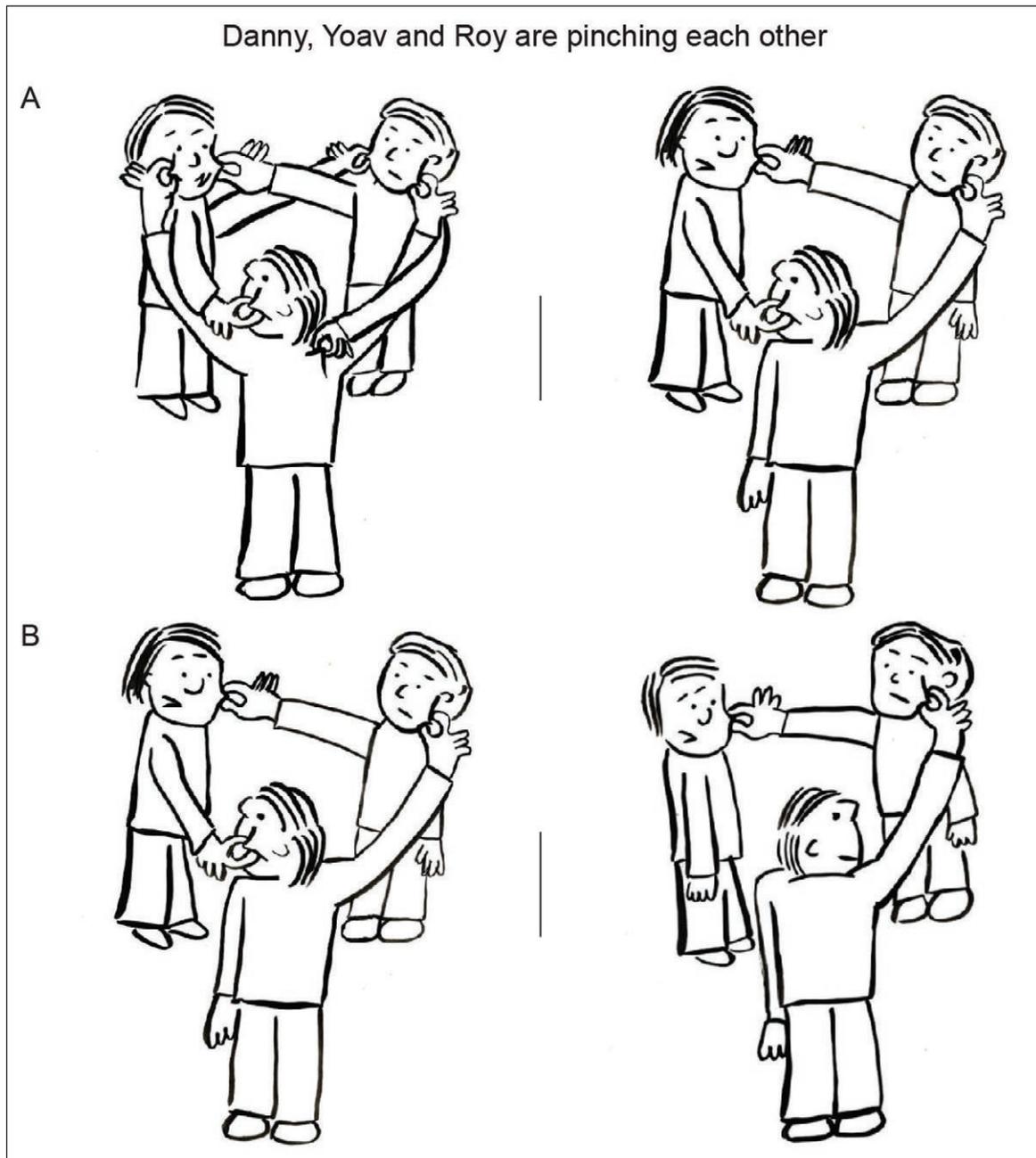


Figure 4: Stimuli from Experiment 1, examples of forced-choice task: main trial S6 vs S3 (A) and control trial S3 vs S2 (B) (translated from Hebrew).

In addition, we included twenty-three filler items which also contained two drawings. These drawings were accompanied by non-reciprocal sentences (e.g. the Hebrew correlate of ‘Owen, Luke and Andrew are painting themselves’). The task was presented as a pen-and-paper questionnaire.

3.2.3 Results

The results of Experiment 1 demonstrated differences between items. For the main test item (S6 vs. S3), the preference for S3 ranged from 10% to 67%. Importantly, for most verbs, a substantial proportion of participants preferred S3 over S6. These results are summarized in Table 1.

For the control item (S3 vs S2), the vast majority of participants preferred S3 over S2 (see Table A2 in Appendix A).

Table 1: Experiment 1 results main test item (*S6* vs *S3*).

Verb (translated from Hebrew)	% preference for <i>S3</i>
Hug (control verb)	10.2
Hit	18.8
Paint	24.5
Pinch	26.5
Point at	29.2
Shake	34.7
Stab	38.8
Apply make-up	40.8
Clean	40.8
Towel	40.8
Give a speech (control verb)	42.9
Scrape	42.9
Comb	66.7

3.3 Discussion

From the pretest, we learned that there are verbs that show a preference for a one-patient situation over a two-patient situation, as expected. We then examined reciprocal sentences that contain those verbs that showed clear one-patient preferences in the pretest. The main part of Experiment 1 tested our expectation that the interpretation of such sentences would be weaker than what the SMH predicts.

The results of Experiment 1 showed a large variability in preferences between *S3* and *S6*. Crucially, for eight out of the thirteen reciprocal sentences that were tested, more than one third of the participants preferred *S3* over *S6*. This result is unexpected by the SMH. The SMH predicts that the meaning of any reciprocal sentence is the strongest one that is consistent with its context. The *S6* situation was explicitly presented to the participants in the forced choice task, hence this situation was possible in the context of the reciprocal sentence.¹⁴ Since this situation is consistent with SR, and SR is the strongest reciprocal meaning, the SMH predicts SR to be the *only* reading of the reciprocal expression. The *S3* situation is inconsistent with this SR meaning, hence the SMH expects 100 percent preference for *S6* over *S3*. This expectation was not borne out: many subjects preferred *S3* to *S6* as the best situation for the reciprocal sentence. It is difficult for the SMH to explain these preferences. Our hypothesis is that these responses are due to the fact that the tested verbs have a typicality preference for one-patient situations. We hypothesize that from the preference for one-patient situations over two-patient situation one can extrapolate a preference for reciprocal situations with only one-patient relations – hence the common preference for *S3* over *S6*. Note that the uniform preference of *S3* over *S2* in the control trials suggests that when typicality does not play a role (there is no difference in patient cardinality between these situations) the preferred situation for the sentence is the core situation (*S3*) that the MTH selects. This conclusion is strengthened by the results of Experiment 2.

¹⁴ An anonymous reviewer points out that illustrations might be discarded as physically unrealistic. If that were the case in Experiment 1, it might in principle salvage the SMH, since those participants who preferred *S3* might have done so because they considered the illustration representing *S6* as impossible. However, if that was the case, we would expect that in a truth-value judgement task with the same verbs, participants would not accept the corresponding reciprocal sentences in *S6*. As we will see, in Experiment 2 the vast majority of participants did accept the parallel reciprocal sentences in Dutch in *S6* situations. Thus, we conclude that the possibility that the reviewer suggests was also not the case with the *S3* preference data in Experiment 1.

4 Experiment 2: Testing the MTH

Experiment 2 tested the Maximal Typicality Hypothesis as a remedy to the problems we encountered for the SMH. The experiment contained two parts. Part 1 tested typicality differences between verb concepts, specifically the preference between one-patient situations vs. two-patient situations. Part 2 tested reciprocal interpretations, specifically whether reciprocal sentences are accepted in *S6* and *S3* situations. When a speaker prefers a one-patient situation for a given verb, the MTH expects the core situation for the reciprocal sentence to be *S3*, with *S6* as another possible situation. Conversely, with verbs that do not show a preference for one-patient situations, the core reciprocal situation is expected to be *S6*. Accordingly, for a representative sample of verbs, the MTH expects a correlation between a verb's *preference* for one-patient situations and *acceptance* of *S3* situations as possible for reciprocal sentences with that verb. Testing this hypothesized correlation was the main goal of Experiment 2.

Part 1 of Experiment 2 was a preference task similar to the pretest of Experiment 1, but included verb concepts that were expected to show a wider range of patient cardinality preferences. This made it possible to test the MTH.¹⁵ In order to study different kinds of verbs within one experiment, both parts of Experiment 2 used a different visual presentation of the stimuli than in Experiment 1: schematic presentation of situations as in Figure 1 and Figure 5 (below).

Part 2 of Experiment 2 made use of a truth-value judgement task for testing interpretations of reciprocal sentences. The MTH makes predictions about acceptability of reciprocal sentences in situations like *S3* and *S6*. Unlike the SMH, the MTH expects there to be reciprocal sentences that are acceptable in both *S3* and *S6*. Thus, while the considerable preference rates for *S3* in the preference task of Experiment 1 are evidence against the SMH, we need truth-value judgements in order to support the MTH. As the strongest test for the MTH, we performed a correlation analysis to test the fine-grained relationships that the MTH predicts, between the acceptability of *S3* in part 2 and the preference for one-patient situations in part 1.

For convenience of presentation, we categorize the verbs used in Experiment 2 into three types, based on their expected patient cardinality preferences:

Type 1 – Neutral verbs: Verbs for which we expected no preference between instances with different patient cardinality (e.g. *know*).

Type 2 – One-patient-preference verbs: Verbs for which we expected a preference for situations with one patient, even though we expected two-patient situations to also be categorized as instances of the verb concept (e.g. *pinch*). This group of verbs was in the focus of Experiment 1.

Type 3 – Strong one-patient-preference verbs: Verbs for which we expected a high preference for one-patient situations, because we expected two-patient situations to not be categorized as instances of the verb concept at all (e.g. *bite*).¹⁶

¹⁵ This is in contrast to Experiment 1, where the goal was to test the SMH using concepts with a one-patient preference. In Experiment 1, this goal dictated a selection of verbs that showed low distribution of such preferences ($M = .79$, $SD = .08$ for the 11 target verbs that were selected for Experiment 1). The two control verbs (*hug* and *give a speech*), which were intended to add more variation, instead showed unexpected effects (see Table 1). This is likely due to factors other than patient cardinality, for example the difficulty of illustrating those verbs in *S3* and *S6*, and the influence of other typicality parameters such as preference for one *agent* acting on any patient (rather than more agents).

¹⁶ Note that the difference between type 2 and type 3 verbs is not expected to surface when measuring the preference for one-patient situations over two-patient situations (as in part 1 of Experiment 2). However, the difference between the two types is expected to show in truth-value judgments of reciprocal sentences in situations that contain two-patient situations, i.e. *S6* (in part 2 of Experiment 2). Therefore, the two groups are distinguished here, although the difference between them is not measured in the preference task of Part 1.

This intuitive classification was used in the experiment as a means for selecting candidate verbs, as well as for presenting large-scale differences between groups of verbs.

4.1 Pretest

In preparation for Experiment 2, we conducted a pretest in Dutch that measured only patient cardinality preferences. This pretest had several aims. Firstly, it tested whether measuring typicality effects with schematic stimuli is comparable to measuring them with pictorial stimuli as in Experiment 1. Secondly, the pretest tested whether Dutch verbs behave in a comparable way to their Hebrew counterparts in Experiment 1. Thirdly, and most importantly, the results of the pretest were used for selecting the verbs to be tested in the main parts of Experiment 2. In part 1 of the experiment, we aimed to obtain a wide range of cardinality preference values. This would ensure that the eventual correlation analysis with reciprocal interpretations is based on verbs that show diverse typicality effects. To this end, the pretest was used for selecting those verbs that showed the clearest differences in patient cardinality preferences.

The pretest measured patient cardinality preferences for 60 Dutch verbs, 32 of which overlapped with the verbs that were used in the pretest for Experiment 1. As in Experiment 1, we measured whether there was a preference between situations with one patient per agent vs. situations with two patients per agent.

4.1.1 Participants

Twenty-one Utrecht University students (19 female, age $M = 21$) participated for monetary compensation. All participants were native Dutch speakers without dyslexia. Prior to the experiment, all participants signed an informed consent.

4.1.2 Materials

Sixty different Dutch verbs were tested. These included 16 verbs that were a priori assumed to be of type 1 (neutral verbs like *kennen*, ‘know’), 29 verbs that were a priori classified as type 2 (one-patient-preference verbs like *knijpen*, ‘pinch’), and 8 verbs that were a priori classified as type 3 (strong one-patient-preference verbs like *bijten*, ‘bite’). In addition, we included 7 verbs like *fotograferen* (‘photograph’) and *ontmoeten* (‘meet’) (see Table B1 in Appendix B) that were used as filler verbs. These verbs have a typical ‘collective’ interpretation in which an activity is performed on a collection of patients rather than on individual patients. These verbs were expected to show some preference for two patients over one. They were added to achieve an optimal balance in typicality preferences.

For each verb, we included one experimental pair of schematic representations, reflecting the choice between two instances of that verb – one with one patient and one with two patients. Each schema included three individuals, which were represented by three proper names, and one or more arrows between them, reflecting either one or two actions between the three individuals. An example of an experimental pair appears in the top row of Figure 5.

In addition, for each verb we included five filler pairs in order to control for visual complexity, by adding all possible arrow combinations, and response bias, by alternating the placement of the different configurations. The filler items differed from the experimental pair in the number of arrows (Figure 5).

The experimental pairs (1 per verb) and filler pairs (5 per verb) for the 60 verbs resulted in a total of 360 trials. The trials were presented in a pseudo-random order with the restriction that no verb or schematic pair would repeat in two consecutive trials. The position of the different schematic representations (on the left or on the right) and the

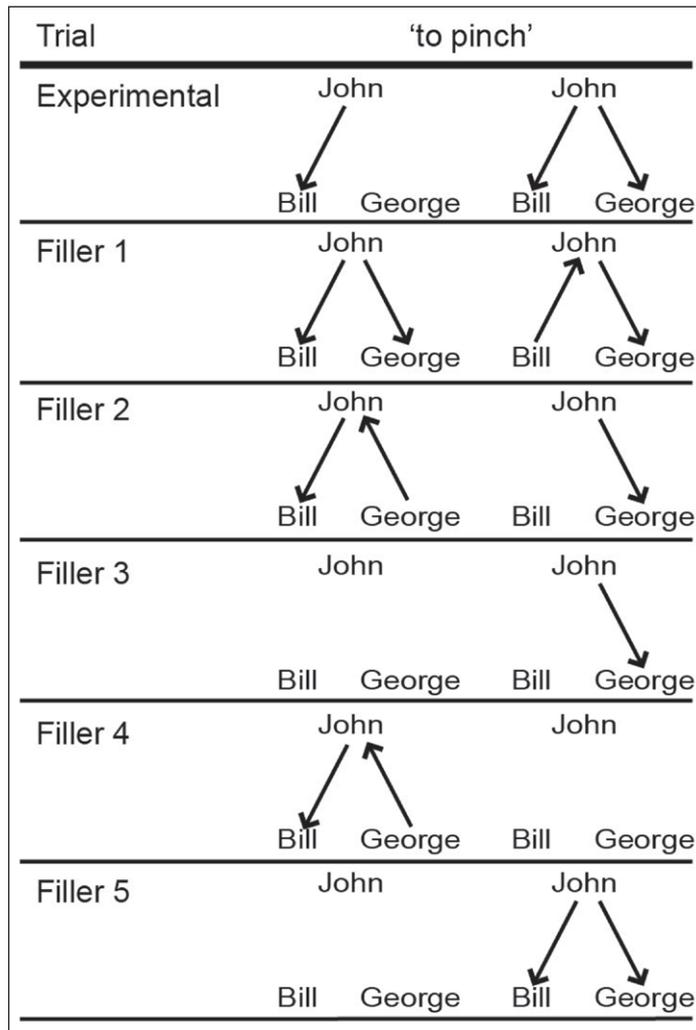


Figure 5: Examples of pairs for *pinch* in the pretest and in part 1 of Experiment 2 (translated from Dutch).

pointing direction of the arrows (to the left or to the right) were counterbalanced over the trials and the verbs.

4.1.3 Procedure

The task was presented in Dutch in a sound-proof booth on a PC using *Presentation* software (Neurobehavioral Systems, Albany, CA). Prior to entering the booth, each participant was instructed verbally about the set-up and on how to interpret the schematic representations. Further instructions were given on the PC monitor, stressing that each schematic representation should be interpreted as representing a situation at one point in time, rather than multiple situations over a time interval. This was clarified in order to make sure that the verbs are interpreted in the same tense as in part 2 of Experiment 2 (see footnote 8). After the instructions, each participant completed six practice trials. Subsequently, participants were given the opportunity to ask for further clarifications, followed by six additional practice trials. No verbs that were used in the practice trials were used in the actual pretest. The pretest itself consisted of two blocks of trials. Each trial started with a fixation cross (500 ms), followed by the presentation of a verb in the top centre of the screen and, below the verb, a pair of schematic representations: one of the six pairs from Figure 5. Participants were instructed to select the schematic representation that best represented the given verb by pressing the left or right arrow

key accordingly, with their dominant hand. The verb and the schematic representations remained visible on the screen for 5000 ms, or until the participant responded.

4.1.4 Analysis

We calculated the proportion of reactions to the test items where a schematic representation of a one-patient situation was selected. We performed a correlation analysis on the patient cardinality data on the 32 Hebrew verbs that were tested in the pretest for Experiment 1, and the corresponding Dutch verbs that were examined in the current pretest.

4.1.5 Results

The results of the pretest are given in Table B1 of Appendix B. These results about Dutch verbs show a significant one-sided positive correlation with the patient cardinality preferences of the corresponding 32 Hebrew verbs that were tested in the pretest for Experiment 1 ($r(32) = .39$, p (one-tailed) = .014), which indicates that the schematic method yields comparable patient cardinality preferences to the pictorial method of Experiment 1. This lends support to our assumption that the two methods measure the same feature, and that the Hebrew and Dutch verbs that we tested have similar meanings.

The results of the pretest were used to select verbs for Experiment 2. The selection procedure is explained in the Materials section of section 4.2.

4.2 Experiment 2 (part 1): Typicality

Part 1 of Experiment 2 is a replication study that measured patient cardinality preferences for a subset of the Dutch verbs that were used in the pretest. This part again measured patient cardinality preferences for the selected subset of verbs. Therefore, we expected to see similar behavior to what we observed in the pretest: verbs that we classified as type 1 verbs were expected to show no preference, those that we classified as type 2 verbs were expected to show a preference for one-patient situations, and those that we classified as type 3 verbs were expected to show the same preference as type 2 verbs, possibly more substantially.

4.2.1 Participants

Eighteen Utrecht University students (15 female, age $M = 21$) participated for monetary compensation. All participants were native Dutch speakers without dyslexia and did not participate in the pretest. Prior to the experiment, all participants signed an informed consent.

4.2.2 Materials

Based on the results of the pretest, we selected 18 verbs that were to be tested regarding their patient cardinality preferences (see Table B2 in Appendix B). The selection process went as follows. Firstly, to minimize confounds coming of syntactic processing we only selected transitive verbs, and ruled out verbs that would require a preposition when used in a reciprocal sentence (e.g. *luisteren naar* 'listen to'). These verbs are marked with an asterisk in Table B1 in Appendix B. Secondly, we selected six verbs from each of the three classes of verbs. Additionally, from the type 1 verbs, we also ruled out verbs where we expected S6 situations to be infelicitous for reasons that are independent of patient typicality. For example, with the verb *horen* ('hear'), participants showed almost an equal preference (.524) for one patient and for two patients. However, in an S6 situation, a person simultaneously hears two other people while they are hearing her. Such a situation might be too noisy for the sentence *they are hearing each other* to

make sense. Other verbs that were ruled out for similar reasons are: *complimenteren* ('compliment'), *noemen* ('name'), and *zien* ('see'). These verbs are marked with double asterisks in Table B1 in Appendix B. After applying this restriction, we selected the six verbs that had the lowest preference for a one-patient situation. From the type 2 verbs (one-patient-preference) as well as the type 3 verbs (strong one-patient-preference), we expected fewer confounds than with type 1 verbs. Therefore, we simply selected the six verbs from each of the two classes that showed the highest preference for a one-patient situation (on the reasons for the distinction between type 2 and type 3, see footnote 13).

The 18 verbs selected for Experiment 2 are the Dutch correlates to the following verbs:

Type 1 (neutral):	<i>envy, know, understand, admire, miss, hate</i>
Type 2 (one-patient-preference):	<i>pinch, hit, caress, stab, shoot, grab</i>
Type 3 (strong one-patient-preference):	<i>kiss, dress, kick, lash out, bite, lick</i>

Similarly to the pretest, for each of the selected 18 verbs we used one experimental pair and five filler pairs to control for visual complexity, by adding all possible arrow combinations, and response bias, by alternating the placement of the different configurations (Figure 5). This resulted in a total of 108 trial items, which were presented in a pseudo-random order with the restriction that no verb or schematic pair would repeat in two consecutive trials. The position of the different schematic representations (on the left or on the right) and the pointing direction of the arrows (to the left or to the right) were counterbalanced over the trials and the verbs.

4.2.3 Procedure

The procedure of the task in part 1 was identical to the procedure of the pretest.

4.2.4 Analysis

We calculated the proportion of reactions to the test items where a schematic representation of a one-patient situation was selected. We then performed a correlation analysis on this patient cardinality data and the patient cardinality data from the same verbs in the pretest, in order to verify that the results were comparable. Next, we performed a generalized linear mixed model (GLMM) logistic regression analysis in SPSS (v.24) on the responses to the test items (Quené and van den Bergh 2008). The responses to the one-patient situation were coded as 1 and to the two-patient situation were coded as 0. The model was built iteratively, starting from a basic model and adding relevant predictors one by one and assessing whether the model accuracy improved. In the final model the fixed part contained Verb type with three levels (1, 2 and 3). The random part of the model contained participants as random effect. The random effect of verb was also tested, but did not add significantly to the model, see Appendix C1 for details on model comparisons.

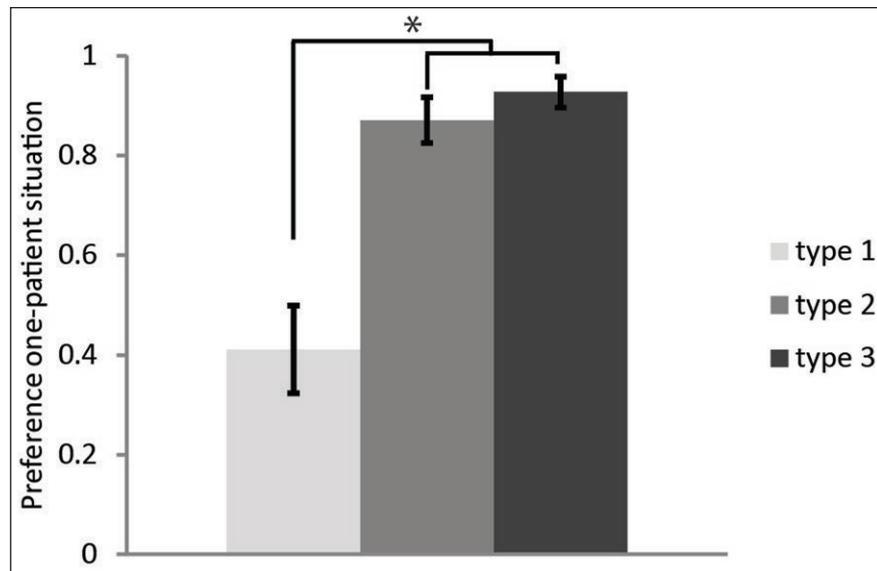
4.2.5 Results

We found a significant one-sided correlation ($r(18) = .57$, p (one-tailed) = .007) between the results for the 18 verbs in part 1 of Experiment 2 and those same 18 verbs in the pretest. This means that as expected, we replicated the pretest for the 18 selected verbs. The results of the one-patient preference for the 18 verbs are presented in Table 2.

For the 18 verbs in part 1 of Experiment 2, we tested the effect of Verb type (1, 2 and 3). The GLMM revealed a significant main effect of Verb type ($F(2, 319) = 31.70$, $p < .001$). As expected, the results show strong preference for a one-patient situation

Table 2: Experiment 2 part 1 results of one-patient preference for the tested verbs.

Verb	Type	Preference one-patient
Grab	1	0.94
Shoot	1	0.89
Stab	1	0.88
Caress	1	0.83
Hit	1	0.78
Pinch	1	0.67
Lick	2	1.00
Bite	2	1.00
Lash out	2	0.94
Kick	2	0.89
Dress	2	0.89
Kiss	2	0.65
Hate	3	0.67
Miss	3	0.44
Admire	3	0.39
Understand	3	0.39
Know	3	0.33
Envy	3	0.33

**Figure 6:** Experiment 2 part 1 – analysis of preference for a one-patient situation with three types of verbs. Error bars represent standard errors of the mean.

with both type 2 verbs ($M = .87$, $SE = .05$) and type 3 verbs ($M = .93$, $SE = .03$). Pairwise comparisons showed that there was no significant difference between these preferences ($t(319) = 1.38$, $p = .167$). The preference for one-patient situations with type 1 verbs ($M = .41$, $SE = .09$) was significantly lower compared to type 2 verbs ($t(319) = 6.32$, $p < .001$) and type 3 verbs ($t(319) = 6.67$, $p < .001$). See Figure 6, and the further details in Table B2 in Appendix B.

4.3 Experiment 2 (part 2): Reciprocal interpretation

Part 2 of Experiment 2 tested responses to a truth-value judgement task on Dutch reciprocal sentences in different situations. The verbs that were tested were the same verbs that were used in part 1, and the reciprocal situations were presented using schematic representations. The truth-value judgement task was used to measure to what extent a given reciprocal sentence is accepted in situations *S6* and *S3*. The MTH links typicality preferences of verbs to truth-value judgements on reciprocal sentences as follows: *S6* is expected to be the core situation for sentences with verbs that show no patient cardinality preferences or a preference for two-patient situations; *S3* is expected to be the core situation for sentences with verbs that show a preference for one-patient situations. Accordingly, we expected to see substantially higher acceptance rates with *S3* for verbs of type 2 and 3, compared to the same situation with verbs of type 1.

4.3.1 Participants

A total of 25 Utrecht University students participated for monetary compensation (24 female, age $M = 23$). All participants were native Dutch speakers without dyslexia and did not participate in the pretest or part 1 of the experiment. Prior to the experiment all participants signed an informed consent.

4.3.2 Materials

The same 18 verbs from part 1 of Experiment 2 were used, but now in Dutch reciprocal sentences of the form *A, B and C P each other* (where *A, B and C* are proper names and *P* is a verb). The resulting 18 sentences were tested for their interpretation in a truth-value judgement task.

For each sentence, we included two experimental trials – schemas reflecting *S6* and *S3*. As in Experiment 1, those situations illustrate relations between three individuals. In situation *S6*, each individual acts on both other individuals. In situation *S3*, each individual acts on exactly one other individual and is acted on by exactly one other individual. Each schema included three individuals, which were represented by three proper names, and either six arrows (in *S6*) or three arrows (in *S3*) between them. Examples of these experimental trials are in the top two rows of Figure 7.

In addition, for each verb we included one control trial and two filler trials (Figure 7). The filler trials were used to control for visual complexity and frequency of no-responses. The control trials illustrated a situation in which only two individuals act on another individual (*S2*). We added this control item because it was expected to be a typical instance for all verbs, but not the core situation for any reciprocal sentence containing them. Similarly, for sentences with verbs of type 1, situation *S3* is expected to be typical but not the core situation. By adding the control trial *S2*, we aimed to also have typical but non-core situations for sentences with verbs of types 2 and 3. This allowed us to better test the predictions of the proposed MTH, since this principle also makes predictions about acceptability rates for non-core situations.

The experimental trials (2 per verb), control trial (1 per verb) and filler trials (2 per verb) for all 18 sentences resulted in a total of 90 trials. The trials were presented in a pseudo-random order with the restriction that no verb or schematic representation would repeat in two consecutive items. The pointing direction of the arrows (to the left or to the right and up or down) was counterbalanced over the verbs.

4.3.3 Procedure

Part 2 of Experiment 2 consisted of two blocks of trials. As in part 1, the task was presented in Dutch in a sound-proof booth on a PC with *Presentation* software (Neurobehavioral

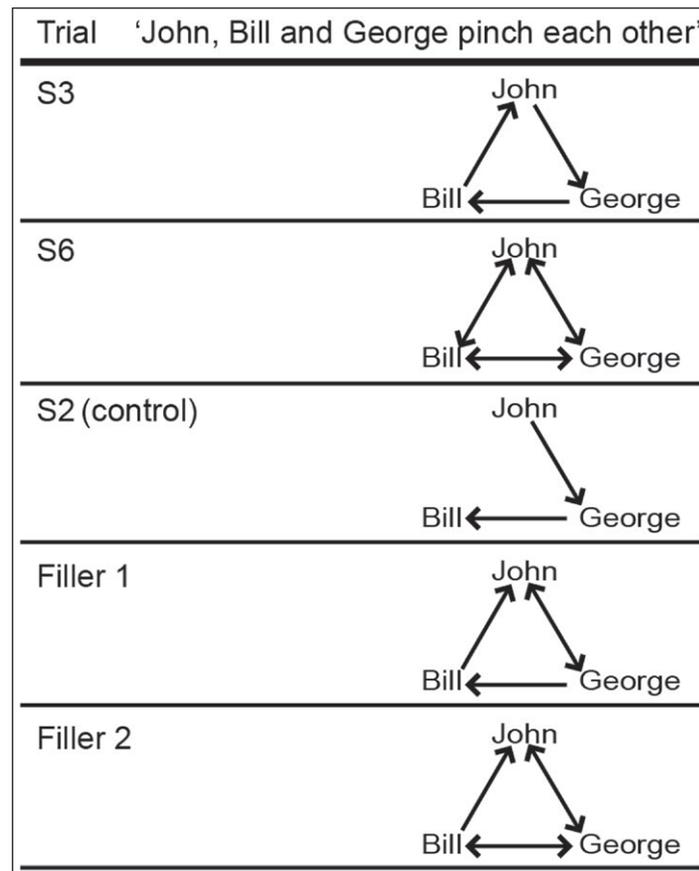


Figure 7: Examples of trials for *John, Bill and George pinch each other* in part 2 of Experiment 2 (translated from Dutch).

Systems, Albany, CA). The instructions on the way schemas were to be interpreted resembled those of part 1, and were similarly followed by practice trials.

Each trial started with a fixation cross (500 ms), followed by the presentation of a reciprocal sentence (e.g. *John, Bill en George knippen elkaar*, 'John, Bill and George pinch each other') at the top of the screen. After 2000 ms, a schematic representation was added to the screen, below the sentence. Participants were instructed to indicate whether the situation that was presented schematically is a possible depiction of the sentence or not, by pressing a green or red button accordingly (right and left arrow key respectively, marked with a sticker), with their dominant hand. The sentence and the schematic representation remained visible on the screen until the participant responded, or for 10000 milliseconds if there was no response.

4.3.4 Analysis

We calculated the proportion of affirmative responses to the control trials (S2) and the experimental trials (S6 and S3), which reflects the acceptability of a given schematic representation of a reciprocal situation as a possible depiction of the given sentence. Further statistical analysis focused on the experimental trials (S6 and S3). We performed a generalized linear mixed model (GLMM) logistic regression analysis in SPSS (v.24) on the responses to the experimental trials. The model was built iteratively, starting from a basic model and adding relevant predictors one by one and assessing whether the model accuracy improved. In the final model the fixed part contained Verb type with three levels (1, 2 and 3) and Trial type with two levels (S6 and S3). The random part of the model

contained participants as random effect. The random effect of verb was also tested, but did not add significantly to the model, see Appendix C2 for details on model comparisons.

4.3.5 Results

Control trial S2: The acceptance of reciprocal sentences in S2 was very low for sentences with all verb types: reciprocal sentences with type 1 verbs ($M = .06, SE = .02$), reciprocal sentences with type 2 verbs ($M = .13, SE = .06$), and reciprocal sentences with type 3 verbs ($M = .12, SE = .05$). For details see Table B2 in Appendix B.

Further analysis of the results of part 2 of Experiment 2 focused on the two experimental trial types that measured acceptability of reciprocal sentences in S6 and S3. There was a significant main effect of Trial type ($F(1, 893) = 10.18, p = .001$) and Verb type ($F(2, 893) = 11.52, p < .001$), as well as a significant interaction between Trial type and Verb type ($F(2, 893) = 30.06, p < .001$). This interaction was further analyzed with pairwise comparisons.

Experimental trial S6: Regarding the acceptance of reciprocal sentences in S6, there were significant differences between all three types of sentences (all p 's $< .001$) (see Figure 8 and Table B2 in Appendix B). The highest acceptability of sentences in S6 was found for those sentences containing type 1 verbs ($M = .98, SE = .01$), followed by those with type 2 verbs ($M = .87, SE = .03$) and those with type 3 verbs ($M = .60, SE = .05$).

Experimental trial S3: Regarding acceptance of reciprocal sentences in S3, the pattern was different (see Figure 8 and Table B2 in Appendix B). There was no significant difference between reciprocal sentences with type 2 verbs ($M = .89, SE = .03$) and reciprocal sentences with type 3 verbs ($M = .84, SE = .04, t(893) = 1.32, p = .189$). Both types of sentences showed high levels of acceptability in S3. The acceptability of reciprocal sentences with type 1 verbs in S3 ($M = .53, SE = .065$) was significantly lower compared to sentences with type 2 verbs ($t(893) = 6.74, p < .001$) and with type 3 verbs ($t(893) = 5.76, p < .001$).

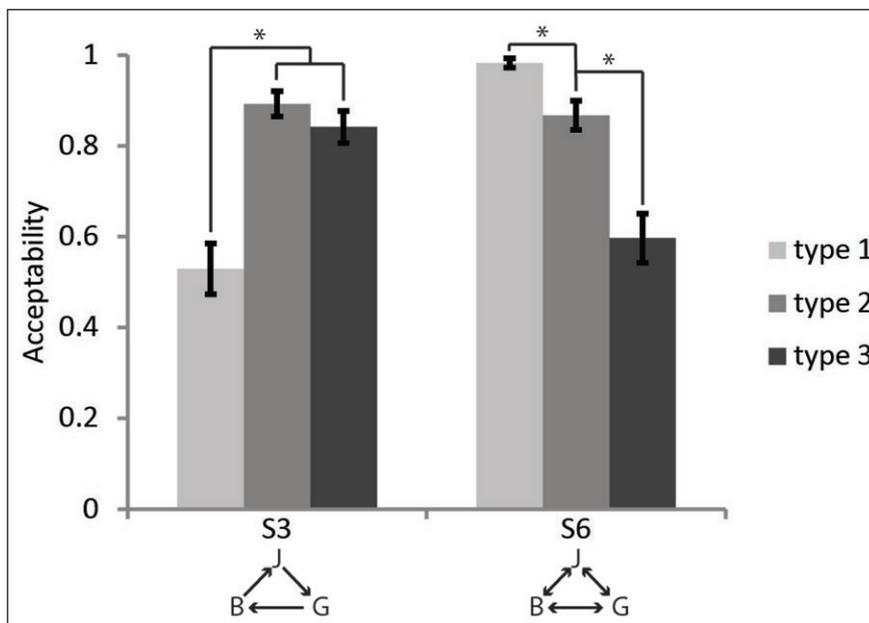


Figure 8: Experiment 2 part 2 – item analysis of acceptance rate for sentences in S6 and S3 with 3 types of verbs. Error bars represent standard errors of the mean.

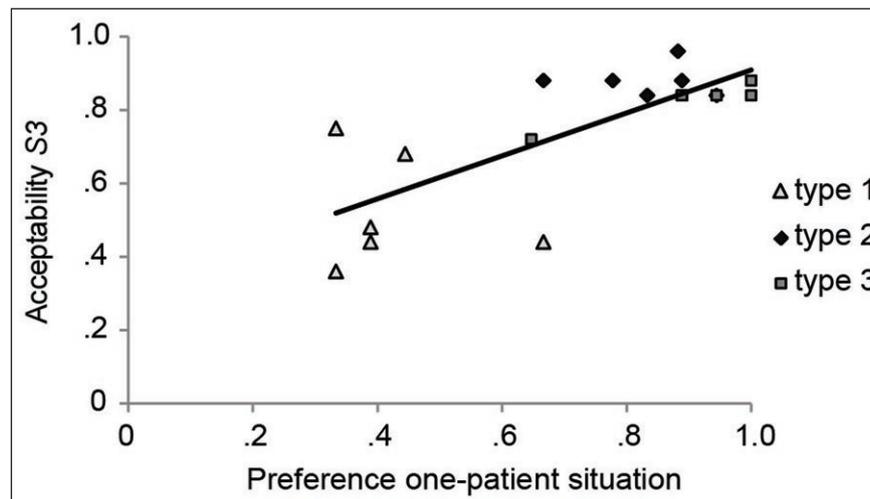


Figure 9: Relation between preferences for one-patient situations in part 1 of Experiment 2, and acceptance of reciprocal sentences in S3 in part 2 of Experiment 2.

4.4 Correlation between results of part 1 and part 2

As expected, the results from part 1 showed a considerable variability in the patient cardinality preferences of different verbs ($M = .72$, $SD = .24$, Table B2 in Appendix B). The MTH predicts that the preference for one-patient situations (as measured in part 1) correlates with the acceptability of reciprocal sentences in S3 (as measured in part 2), explaining the interpretations of reciprocal sentences with type 2 verbs. This prediction was borne out: we found a significant one-sided positive correlation ($r(18) = .76$, $p < .001$, see Figure 9).

4.5 Discussion

The main aim of Experiment 2 was to test the predictions of the proposed Maximal Typicality Hypothesis: whether the core situation described by a reciprocal sentence is maximal among those situations that are most typical for the verb concept in the reciprocal's scope. To that end, part 1 measured typicality effects for 18 different verb concepts, which were selected based on their patient cardinality preferences as measured in a pretest. Part 2 measured whether reciprocal sentences containing those 18 verbs were accepted in S6 and/or S3 (i.e. the two experimental trials). The test for the MTH was in the observed differences between verb types, and in the more fine-grained correlation analysis between the two parts of Experiment 2.

4.5.1 Main results: The core situation

In part 1 of the experiment we found that different verbs show different patient cardinality preferences. As expected, those verbs that we a priori classified as type 1 (neutral) showed no clear preference for an instance with one patient vs. two patients. By contrast, those verbs that we called type 2 (one-patient-preference) and type 3 (strong one-patient-preference) showed a clear preference for instances with one patient over instances with two patients. From these patient cardinality preferences, we extrapolate typicality preferences for reciprocal situations. For type 1 verbs, we extrapolate that there is no difference in typicality between reciprocal situations that differ in terms of patient cardinality, while for verbs of type 2 and 3 we extrapolate that reciprocal situations with one patient per agent are more typical than those with two patients per agent. If, as extrapolated, all situations for type 1 verbs are of equal typicality, then the MTH predicts that the core situation for a reciprocal sentence with a type 1 verb is S6: the maximal

situation among all situations. Thus, for reciprocal sentences with type 1 verbs the MTH expects the acceptability in *S6* in part 2 of the experiment to be high, and substantially higher than in *S3*. For sentences with type 2 and type 3 verbs, the MTH predicts the core situation to be *S3*: the maximal among those situations that have no more than one patient per agent. Thus, the MTH predicts the acceptability of these sentences to be high in *S3*, and substantially higher than the acceptability of reciprocal sentences with type 1 verbs in *S3* situations.

The results of part 2 show that reciprocal sentences with type 1 verbs are very often accepted in *S6* (98% of the time), and substantially more so than in *S3* (53%, $t(893) = 8.22, p < .001$). Our findings also indicate that sentences with type 3 verbs are highly acceptable in *S3* (83%), significantly more so than sentences with type 1 verbs. These findings are expected by the MTH based on the typicality preferences observed in Part 1. The same results are also explained by the SMH, provided that we assume that type 3 verbs like *bite* are judged as physically impossible in situations with two patients per agent (an assumption that we did not test directly in our experiments).

However, when it comes to type 2 verbs, our findings offer substantial support for the MTH over the SMH. Reciprocal sentences with such verbs showed high acceptability rates in *S3* (88%), comparable to type 3 verbs, and significantly more so than sentences with type 1 verbs. This is as expected by the MTH, while the SR meaning that the SMH predicts for these sentences does not hold in *S3*.¹⁷

The success of the MTH in predicting the acceptability pattern for reciprocal sentences in *S3* is further supported by the significant correlation that was found between this acceptability and preferences for one-patient situations as measured in part 1. In this way, the MTH not only explains the overall acceptability of reciprocal sentences with type 2 verbs in *S3*, but gives a fine-grained prediction for the interpretation of other reciprocal sentences in *S3* situations.

4.5.2 Additional results: Non-core situations

So far, we have only discussed the acceptability of reciprocal sentences in the core situation. That is, the high acceptability of sentences with type 1 verbs in *S6*, and the high acceptability of sentences with type 2 and type 3 verbs in *S3*. In addition to that, however, our data provide information about the acceptability of these sentences in non-core situations.

Firstly, one important aspect of our results are the observed acceptability rates of sentences with type 2 and type 3 verbs in *S6*. For both kinds of sentences, the MTH predicts the same core situation: *S3*. However, as we might expect, the acceptability of reciprocal sentences with type 3 verbs (e.g. *bite*) in *S6* is significantly lower than that of sentences with type 2 verbs (e.g. *pinch*) ($t(893) = -5.14, p < .001$). Sentences with type 2 verbs were in fact equally acceptable in *S6* and *S3* ($t(893) = 0.69, p = .492$). Potential differences between type 2 verbs and type 3 verbs with respect to two-patient situations like *S6* were not measured in the preference tasks of part 1. Thus, the difference in interpretations between sentences with type 2 verbs and type 3 verbs calls for a separate explanation. We hypothesize that the high preference for one-patient situations with such verbs (Experiment 2 part 1) has an additional factor with type 3 verbs compared to type 2 verbs. In the case of type 3 verbs, we believe that for many participants, the one-patient preference reflects not merely a choice between two possible situations, but a preference of a possible situation (with one patient) over an impossible, or inconceivable, situation

¹⁷ Note that with type 2 verbs, the acceptability rates of reciprocal sentences in *S6* situations were as high as in *S3* situations. This means that SR cannot be ruled out, and accordingly the SMH expects *S3* not to be accepted (see also footnote 14 above).

(with two patients). For type 2 verbs, one-patient situations are uniformly preferred over two-patient situations, but the latter are likely to be accepted as instances of the verb concept, as witnessed by the high acceptability rates of sentences with such verbs in *S6*. Such a difference between type 2 and type 3 verbs could not be measured in the forced choice task of part 1, but it would immediately affect acceptability judgements of reciprocal sentences in *S6* in part 2. If a participant thinks that a two-patient situation is not possible for type 3 verbs like *bite*, she will judge *S6* as strictly speaking impossible for a reciprocal sentence with such verbs. For type 2 verbs, two-patient situations put less strain on the imagination of the participants. Accordingly, *S6* situations are usually accepted for reciprocal sentences with type 2 verbs.¹⁸ This pattern with *S6* situations and type 2 and 3 verbs does not come as a surprise: it fully agrees with our *a priori* classification of type 2 and type 3 verbs, which was based on mere introspection, and it is consistent with the results of Experiment 1 on type 2 verbs and the assumption of SMH-based accounts on type 3 verbs. Therefore, we believe that simple experimental measures can distinguish type 2 verbs from type 3 verbs: e.g. asking participants to mark possible situations, rather than to choose between them. Running such an experiment would be unproblematic, if relevant for further research.

Another question on non-core situations concerns the status of *S3* for sentences with type 1 verbs. For such verbs, *S3* situations are properly contained in their core situation: *S6*. As expected, reciprocal sentences with those verbs are fully acceptable in *S6*. The SMH accounts for this fact using the strong reciprocity operator, which expects these sentences to be downright unacceptable in all situations that are properly contained by *S6*. The acceptability rates of sentences with type 1 verbs in *S3* in our experiment ($M = .60$, $SE = .06$) go against this prediction of the SMH, especially when compared to the downright unacceptability of the same sentences in *S2*. This problem for the SMH appears because it makes standard binary (true/false) predictions about reciprocal sentences. Accordingly, the SMH expects decisive judgements on sentences in all situations. Unlike the SMH, the MTH does not use such absolute terms for acceptability of reciprocal sentences in situations that are properly contained in the core situation. The MTH only expects this acceptability to be lower than the acceptability in the core situation. Thus, sentences with type 1 verbs are expected to show decreasing acceptability in *S2* and *S3* situations compared to *S6*. Similarly, sentences with type 2 and 3 verbs are expected to show lower acceptability in *S2* than in *S3*. Our acceptability results on *S2* and *S3* are consistent with these predictions.

The acceptability rates in *S3* are also relevant for evaluating another influential analysis of reciprocals. According to Langendoen's (1978) operator of Weak Reciprocity, a sentence like *the girls know each other* should mean "every girl knows another girl and is known by another girl", which is true in the *S3* situation. However, only 36% of the participants in Experiment 2 accepted the Dutch sentence with the verb "know". More generally, our results show a disadvantage for Langendoen's account for reciprocals with type 1 verbs. Reciprocal sentences with these verbs are significantly less acceptable in *S3* than in *S6*, whereas Langendoen's Weak Reciprocity is equally satisfied by both situations. Thus, some selection principle should be superimposed on any analysis of reciprocals, and explain the marginality of type 1 verbs in *S3* situations, as opposed to other verbs. This is in line with the main conclusions of the present paper.¹⁹

¹⁸ This is consistent with the results of the forced-choice task of Experiment 1, where *S6* was often preferred to *S3* for type 2 verbs, despite the high typicality of one-patient situations.

¹⁹ A similar point is made by Dalrymple et al. (1998: 165), on the basis of their intuitive judgements about the sentence "House of Commons etiquette requires legislators to address only the speaker of the House and refer to each other indirectly".

5 Conclusion

Previous work has predicted that sentences like *the boys know each other* and *the boys pinch each other* are both true only if each of the boys knows/pinches all of the other boys. Our experiments show that this reading is indeed preferred for verbs like *know*. However, with verbs like *pinch*, reciprocals show a weaker meaning. This challenge for previous accounts is addressed by considering that situations where one boy pinches two boys simultaneously are quite atypical. According to our proposed account, when speakers observe this kind of non-typicality, it boosts weaker interpretations of reciprocal sentences. This tendency explains why sentences like *John, Bill and George pinch each other* are as acceptable when each boy pinches only one other boy, as when each boy (atypically) pinches two other boys. Furthermore, our experiments showed that the more atypical the interpretation of a reciprocal sentence, the higher the tendency of speakers to accept weaker interpretations. These findings were formally described using a new principle, the *Maximal Typicality Hypothesis*, which specifies a meaning for a reciprocal expression based on the lexical preferences of the predicate it combines with.

Abbreviations

ACC = accusative; MASC = masculine; PRES = present; SG = singular.

Additional files

The additional files for this article can be found as follows:

- **Appendix A:** Results Experiment 1. DOI: <https://doi.org/10.5334/gjgl.180.s1>
- **Appendix B:** Results Experiment 2. DOI: <https://doi.org/10.5334/gjgl.180.s1>

Acknowledgements

Eva B. Poortman and Marijn E. Struiksmal contributed equally to this paper.

The work by the first, second and fifth author was partially supported by a VICI grant 277-80-002 of the Netherlands Organisation for Scientific Research (NWO). Work by the first and last author was also partially funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 742204). Work by the fifth author was also partially supported by an NWO grant “Reciprocal Expressions and Relational Processes in Language” (2007/8). The work by the third author was partially supported by the Israel Science Foundation (grant 2005231). Figures 2, 3, 4 and all drawings used in Experiment 1 are by Ruth Noy Shapira. We are grateful to Michal Biran for help with Experiment 1, and to Gideon Keren and Alda Mari for extensive comments on a previous draft of this paper.

Competing Interests

The authors have no competing interests to declare.

References

- Beck, Sigrid. 2001. Reciprocals are definites. *Natural Language Semantics* 9. 69–138. DOI: <https://doi.org/10.1023/A:1012203407127>
- Dalrymple, Mary, Makoto Kanazawa, Yookyung Kim, Sam Mchombo & Stanley Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21. 159–210. DOI: <https://doi.org/10.1023/A:1005330227480>
- Dotlačil, Jakub. 2013. Reciprocals distribute over information states. *Journal of Semantics* 30(4). 423–477. DOI: <https://doi.org/10.1093/jos/ffs016>

- Heim, Irene, Howard Lasnik & Robert May. 1991. Reciprocity and plurality. *Linguistic Inquiry* 22. 63–101.
- Kerem, Nir, Naama Friedmann & Yoad Winter. 2009. Typicality effects and the logic of reciprocity. In Ed Cormany, Satoshi Ito & David Lutz (eds.), *Proceedings of Semantics and Linguistic Theory (SALT)* 19. 257–274. eLanguage. DOI: <https://doi.org/10.3765/salt.v19i0.2537>
- Langendoen, D. Terence. 1978. The logic of reciprocity. *Linguistic Inquiry* 9(2). 177–197.
- Lee, Choonkyu. 2017. Typicality knowledge and the interpretation of adjectives. In James A. Hampton & Yoad Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology* (Language, Cognition, and Mind) 3. 123–138. Springer, Cham. DOI: https://doi.org/10.1007/978-3-319-45977-6_5
- Margolis, Eric & Stephen Laurence. 1999. Concepts and cognitive science. In Eric Margolis & Stephen Laurence (eds.), *Concepts: Core readings*, 3–82. Cambridge, MA: MIT Press.
- Mari, Alda. 2013. *Each Other*, asymmetry and reasonable futures. *Journal of Semantics* 31(2). 209–261. DOI: <https://doi.org/10.1093/jos/fft003>
- Murray, Sarah E. 2007. Dynamics of reflexivity and reciprocity. In Maria Aloni, Paul Dekker & Floris Roelofsen (eds.), *Proceedings of the Sixteenth Amsterdam Colloquium*, 157–162. Amsterdam: Institute for Logic, Language, and Computation.
- Murray, Sarah E. 2008. Reflexivity and reciprocity with(out) underspecification. In Atle Grønn (ed.), *Proceedings of Sinn und Bedeutung* 12. 455–469. Oslo: ILOS.
- Poortman, Eva B. 2014. Between intersective and ‘split’ interpretations of predicate conjunction: The role of typicality. In Judith Degen, Michael Franke & Noah Goodman (eds.), *Proceedings of the formal & experimental pragmatics workshop*, 36–42. Tübingen.
- Poortman, Eva B. 2017a. *Concepts and plural predication: The effects of conceptual knowledge on the interpretation of reciprocal and conjunctive plural constructions*. Utrecht: Utrecht University dissertation.
- Poortman, Eva B. 2017b. Concept typicality and the interpretation of plural predicate conjunction. In James A. Hampton & Yoad Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology* (Language, Cognition, and Mind) 3. 139–162. Springer. DOI: https://doi.org/10.1007/978-3-319-45977-6_6
- Poortman, Eva B. & Liina Pyllkkänen. 2016. Adjective conjunction as a window into the LATL’s contribution to conceptual combination. *Brain and Language* 160. 50–60. DOI: <https://doi.org/10.1016/j.bandl.2016.07.006>
- Quené, Hugo & Huub van den Bergh. 2008. Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language* 59(4). 413–425. DOI: <https://doi.org/10.1016/j.jml.2008.02.002>
- Roberts, Craige. 1987. *Modal subordination, anaphora, and distributivity*. Amherst, MA: University of Massachusetts Amherst dissertation.
- Rosch, Eleanor. 1973. On the internal structure of perceptual and semantic categories. In Timothy E. Moore (ed.), *Cognitive development and the acquisition of language*, 111–144. New York: Academic Press.
- Rosch, Eleanor & Carolyn B. Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7. 573–605. DOI: [https://doi.org/10.1016/0010-0285\(75\)90024-9](https://doi.org/10.1016/0010-0285(75)90024-9)
- Sabato, Sivan & Yoad Winter. 2012. Relational domains and the interpretation of reciprocals. *Linguistics and Philosophy* 25(3). 191–241. DOI: <https://doi.org/10.1007/s10988-012-9117-x>

- Scontras, Gregory & Noah D. Goodman. 2017. Resolving uncertainty in plural predication. *Cognition* 168. 294–311. DOI: <https://doi.org/10.1016/j.cognition.2017.07.002>
- Smith, Edward E., Edward J. Shoben & Lance J. Rips. 1974. Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review* 81(3). 214–241. DOI: <https://doi.org/10.1037/h0036351>
- Winter, Yoad. 2017. Critical typicality: truth judgements and compositionality with plurals and other gradable concepts. In James A. Hampton & Yoad Winter (eds.), *Compositionality and Concepts in Linguistics and Psychology* (Language, Cognition, and Mind) 3. 163–190. Springer. DOI: https://doi.org/10.1007/978-3-319-45977-6_7

How to cite this article: Poortman, Eva B, Marijn E. Struiksmā, Nir Kerem, Naama Friedmann and Yoad Winter. 2018. Reciprocal expressions and the Maximal Typicality Hypothesis. *Glossa: a journal of general linguistics* 3(1): 18.1–30, DOI: <https://doi.org/10.5334/gjgl.180>

Submitted: 22 June 2016 **Accepted:** 19 September 2017 **Published:** 02 February 2018

Copyright: © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS The Open Access logo, which is a stylized 'a' inside a circle.