

# Working Memory Mechanism in Proportional Quantifier Verification

Marcin Zajenkowski · Jakub Szymanik ·  
Maria Garraffa

Published online: 30 December 2013  
© Springer Science+Business Media New York 2013

**Abstract** The paper explores the cognitive mechanisms involved in the verification of sentences with proportional quantifiers (e.g. “More than half of the dots are blue”). The first study shows that the verification of proportional sentences is more demanding than the verification of sentences such as: “There are seven blue and eight yellow dots”. The second study reveals that both types of sentences are correlated with memory storage, however, only proportional sentences are associated with the cognitive control. This result suggests that the cognitive mechanism underlying the verification of proportional quantifiers is crucially related to the integration process, in which an individual has to compare in memory the cardinalities of two sets. In the third study we find that the numerical distance between two cardinalities that must be compared significantly influences the verification time and accuracy. The results of our studies are discussed in the broader context of processing complex sentences.

**Keywords** Quantifiers · Computational complexity · Approximate number sense · Working memory · Cognitive control

## Introduction

Many experimental studies has been recently devoted to the processing of natural language quantifiers (e.g. “all”, “most”). Researchers have attempted to determine cognitive functions

---

M. Zajenkowski (✉)  
Faculty of Psychology, University of Warsaw, Stawki 5/7, 00-183 Warsaw, Poland  
e-mail: zajenkowski@psych.uw.edu.pl

J. Szymanik  
Institute for Logic, Language and Computation, University of Amsterdam,  
Amsterdam, The Netherlands

M. Garraffa  
School of Education, Communication and Language Sciences,  
Newcastle University, Newcastle upon Tyne, UK

involved in understanding sentences with quantifiers. For example, [McMillan et al. \(2005\)](#) studied the pattern of neuroanatomical recruitment, while subjects were judging the truth-value of statements containing natural language quantifiers. The authors considered two standard types of quantifiers: definable in first-order logic (e.g., “all”, “some”, “at least 3”), and non-definable in first-order logic, so-called higher-order (e.g., “an even number of”, “more than half”). Roughly speaking, first order quantifiers can be expressed in terms of object quantification, while higher-order quantifiers demand talking about sets and their cardinalities (see, e.g., [Ebbinghaus et al. 1996](#)). [McMillan et al. \(2005\)](#) concluded that all quantifiers recruit the right inferior parietal cortex, which is associated with numerosity, but only higher-order quantifiers recruit the prefrontal cortex, which is associated with executive resources, like working memory (WM). [Szymanik \(2007, 2009\)](#) proposed that the cognitive difficulty of quantifier processing might be assessed on the basis of the complexity of the corresponding minimal automata that could handle the computational task. [Szymanik and Zajenkowski \(2009, 2010a,b\)](#) confirmed this hypothesis in a series of empirical studies by comparing the processing of various classes of quantifiers with respect to their computational complexity. The authors concluded that proportional quantifiers (“more than half” and “less than half”) are the hardest to verify and engage working memory to the highest degree. Actually, it might be shown that those quantifiers require an algorithmic recognition mechanism with unbounded internal memory ([Van Benthem 1986](#)). During computation, the sizes of two sets need to be compared, e.g., by using an algorithm that may be simulated by a push-down automaton (PDA)<sup>1</sup> For instance, in order to verify the sentence “More than half of the cars are red”, one has to count and hold in memory the number of red cars and then compare it with the total number of cars. No such memorization/comparison is necessary when processing other quantifiers, for example “All cars are red”. This theory seems to be cognitively plausible, according to behavioral ([Szymanik and Zajenkowski 2010a,b](#)), clinical ([Zajenkowski et al. 2011](#)), neuroimaging ([McMillan et al. 2005](#)), and neuropsychological ([McMillan et al. 2006](#)) studies.

Working memory is a very broad structure that includes storage systems and executive functions ([Baddeley and Logie 1999](#); [Logie 2011](#)). Some researchers suggest that only the latter might be crucial for complex sentences comprehension ([Just and Carpenter 1992](#)). Similarly, [Szymanik and Zajenkowski \(2011\)](#) argue that an executive mechanism of WM rather than storage is substantial for processing proportional quantifiers. The authors found that the result of a simple short-term memory task was correlated with proportional quantifiers at the same level as with other types of quantifiers. In other words, a larger storage capacity facilitated the processing of various quantifiers to the same extent. Thus, it was hypothesized that executive functions are responsible for the uniqueness of proportional sentence verification.

Since comprehension of proportional quantifiers seems to be more complex than other types of quantified sentences, in the present paper we decided to investigate closely the working memory mechanisms involved in their processing. As we mentioned above, proportional quantifiers require holding in memory numbers of items (cardinalities of) from two sets, which are then compared. In the present studies we wanted to explore this comparison/integration process.

---

<sup>1</sup> [Chomsky \(1957, 1969\)](#) famously proposed a mathematical model of formal grammars to talk about the complexity of the syntactic constructions. His complexity hierarchy classifies grammatical constructions into regular, context-free, context-sensitive, and recursively enumerable. The higher the construction in the hierarchy, the more difficult it is, especially if it potentially engages WM more. The computational model of quantifier verification has been formulated in terms of automata-theory that exactly corresponds to the Chomsky hierarchy: finite-automata recognize regular languages, PDAs recognize context-free languages, linear-bounded non-deterministic Turing machines correspond to the context-sensitive languages, and finally the class of enumerable languages is recognizable by Turing machines (see, e.g., [Hopcroft et al. 2006](#)).

The research on number sense suggest that at least two phenomena might influence the difficulty of proportional quantifiers processing. It has been observed that the intraparietal sulcus implements nonsymbolic (approximate) representations of cardinalities by associating with each numeral an associated normal distribution on a number line (cf. Clark and Grossman 2007). In other words, a display with 10 items is assumed to yield the number representation ‘10’ most often, but frequently it also yields ‘9’ or ‘11’, and less frequently ‘8’ or ‘12’, etc., with the drop off in probability following the Gaussian distribution. Because of this hypothesized representational scheme, the intraparietal sulcus is also referred to as the Approximate Number System (ANS; Dehaene 1999). Given the Gaussian representation of numerical quantities in the ANS, comparison of two numerical quantities is predicted to be easier (i.e., faster and less prone to error) the further the distance between the two numerical quantities (i.e., a *distance effect*). Moreover, for equal numerical distance, performance also decreases with increasing number size (i.e., *size-effect*). These effects are observed even in the experiments where subjects are asked to compare numbers presented in Hindu Arabic notation.

The first question we tried to answer in the present experiments was whether the comparison process requires substantially more cognitive resources than just simple counting and storing the sizes of two sets. We conducted two studies exploring this problem. In the first, we compared the difficulty of proportional quantifiers (e.g. “More than half of the dots are blue”) and sentences, which referred to the two sets presented in the picture (e.g. “There are eight blue and seven yellow dots”). Both types of verbal structures required counting the same numbers of items, however, the first type of sentences needed additional integration/comparison of data. Thus, we predicted that they would be more difficult in terms of time and accuracy.

In the second investigation, we took the individual differences approach and examined how individuals’ working memory resources are related to the performance of sentences from the first experiment. In particular, we tested the subjects’ storage and executive functions separately to see how they contribute to the semantic process. It was previously suggested that these functions are the most substantial aspects of general working memory (Engle et al. 1999). The extra computation required in order to comprehend the meaning of proportional quantifiers would likely be associated with executive resources to a higher degree than other sentences.

In the third study we further explored the comparison process. We hypothesized that the distance between two cardinalities to be compared can significantly influence the verification time. One possible verification strategy for the sentence “More than half of the dots are blue”, consistent with the automata-model, is as follows: compute and store the number of blue dots and the number of non-blue dots; next, compare the two numbers. The bigger the numerical distance between the two numbers, the more precise and faster the judgments. It follows from the Weber law governing the number system, no matter the mental representation of involved cardinalities (Dehaene 1999).

## Study 1

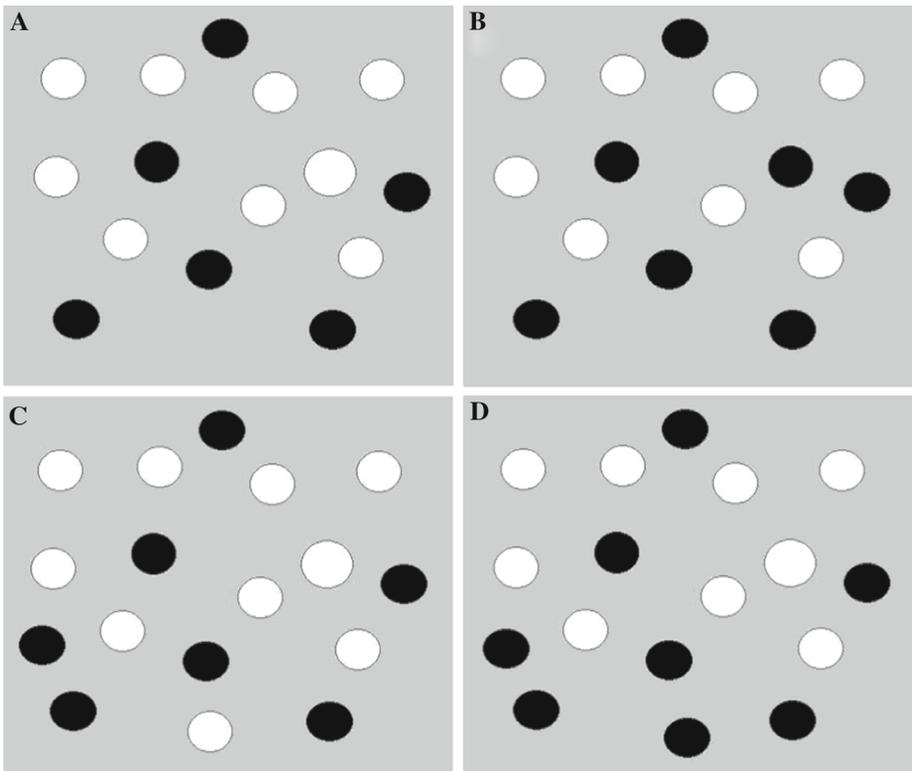
### Participants

A total of 58 subjects took part in the study (35 female and 23 male). Their mean age was 22.87 (SD=2.98). The sample was composed of undergraduate students from the University of Warsaw.

## Measures and Procedure

The task consisted of 32 grammatically simple propositions in Polish that referred to the color of a dot on display. Color pictures of dots accompanied the sentences. For each proposition, the pictures contained 15 or 17 objects in two colors (blue and yellow), as illustrated in Fig. 1. The colors used are suggested for sentence-picture verification tasks (e.g. Halberda et al. 2008). The dots were similar in size (the differences did not exceed one pixel) and were presented in four versions (controlled for the location) in the 15 elements condition and four versions in the 17 elements condition. All the sentences with accompanying pictures were presented randomly. Each quantifier problem resulted in a 18,500 ms event. In the event the proposition and a stimulus array containing randomly distributed cars were presented for 18,000 ms followed by a blank screen for 500 ms. The font of the letters was Times New Roman, size 20. Participants sat in a chair about 90 cm in front of the computer screen. They were asked to decide if the proposition was true of the presented picture. Participants responded by pressing the button marked “P” if true, and the button marked “F” if false. The letters refer to the first letters of the Polish words for “true” and “false”.

There were two factors of interest in this study: sentence type (two levels: proportional and conjunction) and cardinality (two levels: 15 and 17).



**Fig. 1** Four conditions accompanying sentences in study 1 and 2 (picture a and b), and study 3 (A, B, C, D). Picture A illustrates 15 objects of which six were in one color and nine in another; picture b presents 15 objects—seven in one color, eight in another; picture c has 17 items—seven in one color, ten in another; and picture d presents 17 elements—eight in one color, nine in another

Specifically, the proportional sentences were of the form, e.g.:

Więcej niż połowa kropek jest niebieska.

More than half of the dots are blue.

while the conjunction sentences were as follows:

Jest osiem niebieskich i siedem żółtych kropek.

There is eight blue and seven yellow dots.

The second factor, cardinality, was manipulated with the pictures accompanying sentences. They differed in terms of the number of objects (situation A with 15 dots and B with 17 dots are presented in Fig. 1), but not the distance between the cardinalities of two sets of dots (A: 7 vs 8 and B: 8 vs 9), which always equaled one element.

Within each condition subjects had to solve eight trials. Half of them were true. Participants were asked to decide, by pressing a button, whether or not the proposition accurately described the picture. First, we examined the accuracy level. For each participant the number of correct answers was computed. Since there were eight trials within each condition, the scores ranged between 0 and 8. Further, we analyzed reaction time (RT) which was the average from eight responses of each subject.

### Results

First, we examined the accuracy in different linguistic conditions. Because there were two possible responses (correct or incorrect), the general estimation equation model for the binary data was used. There were two predictors: cardinality (15 coded as 0 and 17 coded as 1) and type of sentence (proportional coded as 0 and conjunction coded as 1). The incorrect answer (0) in the dependent variable was the reference level. The analysis revealed a significant interaction effect between predictors (see Table 1). Further analysis conducted separately for each sentence type showed that cardinality was not significantly related with accuracy in both conjunction and proportional judgments. However, when two cardinalities were analyzed separately, it was found that in the situation of 17 objects, sentence type was negatively associated with accuracy. This result suggests, that proportional judgments were performed poorer than conjunction sentences, but only when they were accompanied with a picture displaying a relatively big number of objects.

**Table 1** Results of the logistic regression analysis with sentence type and cardinality as predictors, and accuracy as dependent variable

Parameter	B	SE	95% Wald CI		Wald	df	p
			Lower	Upper			
Sentence type	-.29	.22	-.72	.14	1.73	1	.19
Cardinality	.37	.22	-.08	.81	2.65	1	.10
Sent * card	-.64	.31	-1.25	-.03	4.26	1	.04

**Table 2** Means and standard deviations of reaction time in each condition

	Reaction times		
	15 objects	17 objects	Difference within each sentence type ( $\eta^2$ )
Conjunction	7,385 (1,591)	8,056 (1,857)	0.26
Proportional	8,135 (2,280)	8,141 (2,854)	0.01
Difference between sentence types ( $\eta^2$ )	0.16	0.01	

Next, two-way ANOVA with repeated measures was used to analyze differences in mean RTs. Cardinality (15 or 17) and type of sentence (proportional or conjunction) were two within-subject factors. Table 2 presents means and standard deviations.

The analysis indicated that the main effect of objects number was significant ( $F(1, 57) = 6.12$ ;  $p = 0.016$ ;  $\eta^2 = 0.10$ ), while the effect of sentence type reached the tendency level ( $F(1, 57) = 2.91$ ;  $p = 0.093$ ;  $\eta^2 = 0.05$ ). Additionally, the objects number  $\times$  sentence type interaction was significant ( $F(1, 57) = 7.37$ ;  $p = 0.010$ ;  $\eta^2 = 0.12$ ). The analysis of the interaction effect showed that conjunction sentences accompanied by 15 objects differed significantly ( $p < 0.05$ ) from all other conditions and were processed fastest (see Table 1 for effect sizes).

## Discussion

The study revealed that in the situation in which the number of objects presented in the picture was relatively small, subjects took more time to verify proportional quantifiers than conjunction sentences, but they were equally accurate in both conditions. When the pictures displaying bigger number of items, we observed the speed-accuracy trade off between proportional and conjunction propositions, in the way that proportional judgments were processed faster but poorer than conjunction sentences. The results seem to support the hypothesis derived from previous investigations (Szymanik and Zajenkowski 2011; Zajenkowski et al. 2011) that the crucial cognitive mechanism underlying the complexity of proportional quantifiers is related to the integration process in which an individual has to compare the memorized cardinalities of two sets.

Another difference between two types of sentences is that while the proportional sentences are partitive in nature the corresponding existential sentences are not. In theory, it could relate to the processing differences and therefore in the future it might be worth to compare existential sentences with non-partitive version of our proportional sentences.

Interestingly, there was no difference in subjects' performance on proportional sentences under different semantic conditions. This may suggest that the cognitive difficulty of proportional sentences is related to the integration process and not the counting or storing the particular number of objects. However, the hypothesis about specific cognitive resources involved in the verification of proportional statements should be tested more directly. This was done in the next study in which we examined how specific cognitive resources, related to working memory, contribute to the proportional and conjunction judgment. In particular,

we tested how individuals' level of simple storage and executive function (cognitive control) are associated with semantic processing.

## Study 2

### Participants

Sixty-two subjects participated in the study (40 female and 22 male). Their mean age was 23.00 (SD=2.85). The sample was composed of undergraduate students from the University of Warsaw. Participants from this study did not participate in study 1 and 3.

### Measures and Procedure

In this study we used exactly the same task as in study 1 to measure the performance on proportional and conjunction judgments. Moreover, we tested participants' memory storage and cognitive control.

The memory span task was a computerized version of Sternberg's short-term memory measure (Sternberg 1966). On each trial of the test, the subjects were presented with a random series of different digits, one at a time, for 300 ms, followed by a blank screen and the test digit. Participants had to decide whether the test digit had appeared in the previously displayed string. Sequences of digits of three lengths (four, six, or eight) were repeated eight times each; hence, there were 24 trials overall. The score was the total of correct responses from all conditions (range 0–24).

The cognitive control was measured with the short version of the Attention Networks Test (ANT) designed by Fan et al. (2002). The authors' starting point was the assumption that the attentional system can be divided into three functionally and anatomically independent networks: alerting (allows maintenance of a vigilant and alert state), orienting (responsible for selection of space region to be attended), and executive control (the monitoring and resolution of conflict between expectation, stimulus, and response). In the present study we were focused on the control network as an index of cognitive control. In the ANT task, on each trial, the participant has to decide, by pressing a button, whether a central arrow stimulus (the target) points left or right. The target is flanked by distractor stimuli, and appears above or below a central fixation point. The target stimulus may be preceded by a cue stimulus that either has a general alerting function, or indicates whether the target will appear above or below fixation. Two attributes of the task are manipulated across trials. The first is cue type, which may be absent (central fixation cross only), a central cue (asterisk), or a spatial cue (single asterisk above or below fixation cross). The second attribute is the flanker stimulus type, which may be congruent with the target (arrow points in same direction) or incongruent (arrow points in opposite direction). In each case, two flankers are presented on either side of the target. Each trial consists of the following events: (1) central fixation cross for 400–1,600 ms, (2) cue or no cue for 100 ms, (3) central fixation cross for 400 ms, (4) target until participant responds, and (5) central fixation cross until total trial duration of 4,000 ms has elapsed. The alerting network is calculated by subtracting the RT median of center-cue condition from the RT median of no-cue condition. The orienting index is calculated by subtracting the RT median of the spatial cue conditions from the RT median of the center cue conditions. The executive control index is calculated by subtracting the RT median of the congruent flanking conditions from the RT median of incongruent flanking conditions.

## Results

First, we examined the accuracy in different linguistic conditions with the general estimation equation model for the binary data. There were two predictors: cardinality (15 coded as 0 and 17 coded as 1) and type of sentence (proportional coded as 0 and conjunction coded as 1). The incorrect answer (0) in the dependent variable was the reference level. The analysis revealed a significant interaction effect between predictors (see Table 3). Further analysis of the interaction revealed, that this effect was the same as in study 1. Specifically, the proportional judgments were performed poorer than conjunction sentences, but only when they were accompanied with a picture displaying a relatively big number of objects.

Next, conditions of verbal task were compared with respect to RT using two-way ANOVA with repeated measures, likewise in study 1 (see Table 4 for means and standard deviations). The analysis indicated that the effect of objects number ( $F(1, 61) = 4.13$ ;  $p = 0.047$ ;  $\eta^2 = 0.06$ ) and objects number  $\times$  sentence type interaction ( $F(1, 61) = 6.13$ ;  $p = 0.016$ ;  $\eta^2 = 0.10$ ) were significant, while the effect of sentence type ( $F(1, 61) = 1.18$ ;  $p = 0.280$ ;  $\eta^2 = 0.02$ ) was not significant. The analysis of the interaction effect showed that conjunction sentences presented with 15 objects were processed significantly faster than other propositions. Generally, the pattern of results was similar to the one obtained in the first study.

Next, we correlated all variables with one another (see Table 4). Most importantly, we found that scores on proportional sentences in both conditions were strongly correlated with memory and cognitive control network from the ANT. In the latter case the correlation was negative since the high result on control network indicates delay in inhibiting response to competing stimuli, and hence poor executive functioning. Additionally, the subjects' accuracy on conjunction sentences was moderately and positively related to memory score.

The cognitive control and memory scores were moderately correlated, thus we performed regression analyses to examine the unique contribution of the cognitive functions to the accuracy level of each semantic condition (see Table 5).

The analyses revealed that performance on proportional quantifiers in both semantic conditions was substantially associated with better cognitive control and bigger memory storage, while conjunction sentences were significantly associated only with the latter function.

## Discussion

In the second study we took the individual differences perspective to examine how people with different levels of cognitive control and short-term memory judge truth-values of proportional and conjunction sentences. We found that the task measuring storage function, in the Sternberg's (1966) paradigm, was positively linked to the judgments' accuracy on all

**Table 3** Results of the logistic regression analysis with sentence type and cardinality as predictors, and accuracy as dependent variable

Parameter	B	SE	95% Wald CI		Wald	df	<i>p</i>
			Lower	Upper			
Sentence type	-.39	.22	-.83	.06	2.90	1	.09
Cardinality	.21	.18	-.14	.57	1.35	1	.25
Sent * card	-.57	.26	-1.08	-.05	4.62	1	.03

**Table 4** Descriptive statistics (means and standard deviations) and intercorrelations between all variables from study 2

	Con17 RT	Prop15 RT	Prop17 RT	Con15 ACC	Con17 ACC	Prop15 ACC	Prop17 ACC	Alerting	Orienting	Control	Memory	Mean (SD)
Con15 RT	.78**	.59**	.45**	.02	.16	.07	.07	.08	-.02	-.20	.15	6,934 (1,825)
Con17 RT		.69**	.56**	.15	.24	.28*	.14	.20	-.09	-.08	.07	7,460 (1,939)
Prop15 RT			.80**	.19	.15	.33**	.36**	.02	-.06	-.23	.04	7,447 (2,077)
Prop17 RT				.11	-.03	.30**	.36**	.04	.13	-.16	.06	7,401 (2,203)
Con15 ACC					.61**	.23	.22	-.05	.02	-.20	.30*	7.30 (0.90)
Con17 ACC						.14	.01	.04	-.07	-.16	.30*	7.48 (0.90)
Prop15 ACC							.40**	.20	-.09	-.40*	.37**	6.97 (1.24)
Prop17 ACC								-.14	.03	-.33**	.39**	6.74 (1.42)
Alerting									-.57**	.05	-.12	20.38 (25.12)
Orienting										.23	.02	47.17 (27.10)
Control											-.40*	92.30 (31.23)
Memory												20.60 (2.36)

Con conjunction sentence, Prop proportional sentence, 15 and 17 number of objects, RT reaction time, ACC accuracy, Alerting alerting network (ANT), Orienting orienting network (ANT), control Executive control network (ANT), Memory result from Sternberg memory task

\*  $p < 0.05$   
 \*\*  $p < 0.001$

**Table 5** Results of multiple regression analyses of sentence types accuracy as dependent variables and cognitive functions (control and memory) as predictors

	Con15			Con17			Prop15			Prop17		
	<i>R</i> <sup>2</sup>	$\beta$	<i>p</i>	<i>R</i> <sup>2</sup>	$\beta$	<i>p</i>	<i>R</i> <sup>2</sup>	$\beta$	<i>p</i>	<i>R</i> <sup>2</sup>	$\beta$	<i>p</i>
	.07			.06			.19			.16		
Control		-.10	.46		-.05	.73		-. <b>30</b>	.05		-. <b>21</b>	.09
Memory		<b>.26</b>	.05		<b>.28</b>	.04		<b>.25</b>	.02		<b>.30</b>	.02

*p* < 0.05 are highlighted in bold

kinds of verbal structures used in the study. However, individuals with better controlling capacities as measured by ANT verified only proportional propositions more efficiently. It is worth mentioning that, as some researchers believe, the individual differences in general working memory capacity reflect mainly variation in the ability to control attention (Engle et al. 1999). Indeed, it was found that extreme working memory span groups differ in the executive control network from ANT, but not in the alerting or orienting networks (Redick and Engle 2006).

Theoretically speaking, verification of conjunction sentences require only counting and memorizing two groups of objects. Proportional judgments engage the same functions and additional integration process. The latter should be related to control ability since it deals with competing actions of short-term memory and comparison of information. The results obtained in the current investigation support this hypothesis showing that only proportional judgments involve both functions, related to working memory, to a high degree. Previous findings seem to be in agreement with this conclusion. For instance, in the work of McMillan et al. (2005), described in detail above, the authors concluded that only higher order quantifiers (mainly proportional) activate brain areas associated with executive resources. Similarly, Szymanik and Zajenkowski (2011) found that short-term memory is correlated with proportional judgments almost to the same degree as with other types of quantifier judgments. Moreover, it was shown in other experiment, in which subjects verified sentences with quantifiers while holding arbitrary information in memory, that maintaining irrelevant data resulted in decreased performance but only on proportional statements (Szymanik and Zajenkowski 2011).

Taking all the data into account, we argue that the complexity of the proportional judgments corresponds to high engagement of working memory, and especially to its executive aspect. However, there might be some limitations of this conclusion related to the specific semantic conditions. In particular, under 200 ms time pressure for solving the verification task, it was found that the numerical distance between compared sets might dramatically influence the time and accuracy (e.g. Pietroski et al. 2009). Therefore, this problem was addressed in our next study.

### Study 3

#### Participants

A total of 78 subjects took part in the study (45 female and 33 male). Their mean age was 22.04 (SD=2.25). The sample was composed of undergraduate students from the University of Warsaw. Participants from this study did not participate in study 1 and 2.

## Measures and Procedure

We used a computerized sentence-picture verification task similar to the one used in study 1. It consisted of 32 grammatically simple propositions containing a quantifier that referred to the color of a dot on display. Pictures contained 15 or 17 objects in two colors and were presented simultaneously with a proposition containing a quantifier.

In this study, we tested the verification of proportional quantifiers, e.g. “More than half of the dots are blue”, in different conditions. Specifically, we were interested in two aspects: the total number of dots, and the numerical distance between two compared sets. Therefore, there were two factors of interest in this study: differed in terms of displayed pictures accompanying sentences (see Fig. 1). The first factors were cardinality (two levels: 15 vs. 17 dots) and distance between sets (two levels: one vs. three).

Likewise in study 1, there were eight trials within each condition. Half of them were true. Subjects were asked to decide, by pressing a button, whether or not the proposition was true at the picture. We analyzed mean reaction time (RT) as well as accuracy level (number of correct answers; maximum=8) of each condition.

## Results

First, the accuracy in different linguistic conditions with the general estimation equation model for the binary data was examined. There were two predictors: cardinality (15 coded as 0 and 17 coded as 1) and distance between objects (distance equal three coded as 0 and distance equal 1 coded as 1). The incorrect answer (0) in the dependent variable was the reference level. The analysis revealed that only the distance was related to accuracy (see Table 6). Specifically, the relationship was negative suggesting that the smaller the distance between objects the poorer the accuracy.

Further, two-way ANOVA with repeated measures was used to analyze the differences in RTs. Total number of objects (15 or 17) and distance between compared sets (one element or three elements of difference) were two within-subjects factors. Table 7 presents means and standard deviations for all conditions. The analysis revealed that the main effect of total objects number ( $F(1, 77) = 28.67$ ;  $p < 0.001$ ;  $\eta^2 = 0.27$ ) and distance ( $F(1, 77) = 260.05$ ;  $p < 0.001$ ;  $\eta^2 = 0.77$ ) as well as the objects number  $\times$  distance interaction ( $F(1, 77) = 4.79$ ;  $p = 0.032$ ;  $\eta^2 = 0.06$ ) were significant. Furthermore, we examined the interaction. First, we found that in both distances the processing was slower when there were 17 objects in the picture. However, in case of distance=3 the effect size was much bigger than in distance=1 (see Table 4). Second, there were significant differences between distances within both situations with 15 and 17 objects. In summary, the judgment time increased as follows: distance=3 with 15 objects, distance=3 with 17 objects, distance=1 with 15 objects, and distance=1 with 17 objects.

**Table 6** Results of the logistic regression analysis with distance and cardinality as predictors, and accuracy as dependent variable

Parameter	B	SE	95% Wald CI		Wald	df	<i>p</i>
			Lower	Upper			
Distance	−.91	.26	−1.42	−.41	12.50	1	.00
Cardinality	.26	.19	−.65	.13	1.68	1	.20
Sent * Card	.27	.26	−.24	.78	1.10	1	.29

**Table 7** Means (standard deviations) of verification times (in milliseconds) in each experimental condition

	Reaction time		
	15 objects	17 objects	Difference within the distance ( $\eta^2$ )
Distance = 3	5,981 (2,025)	6,817 (2,433)	0.25
Distance = 1	8,655 (2,346)	8,984 (2,738)	0.06
Difference between the distances ( $\eta^2$ )	0.72	0.63	

## Discussion

In the third study, we compared the difficulty of proportional quantifier processing under different semantic conditions. In particular, we examined how subjects verify proportional sentences accompanied by pictures differing in the number of objects and, most importantly, in the numerical distance of two sets of objects that were compared. We found that the bigger the distance between groups of elements, the better performance (lower RT and higher accuracy). This result was independent of the total number of objects presented in the picture.

Some researchers suggest that people may use an approximation mechanism, rather than algorithmic method, to verify the truth-value of the proportional judgments, especially when the distance between sets is substantial or there is no time for precise judgments (e.g. Pietroski et al. 2009). The question arises whether people used approximation or precise counting in our study. We believe that there are some arguments for the latter interpretation. First, subjects were instructed to take as much time as they needed for precise judgment. Second, in studies such as the one conducted by Pietroski et al. (2009) an experimental procedure is used in which the picture is presented for only 200 ms. In our investigation, mean RTs ranged from approximately six to almost nine seconds and no outliers with extremely short RTs were found. Moreover, the averages in our sample were close to the times obtained in the experimental procedures where precise counting was necessary (e.g., Hackl 2009).

An algorithm based on precise counting has to take into account all objects presented at the picture regardless of the distance between sets. From this point of view, it does not matter whether we want to verify a sentence against a picture containing seven vs. eight objects or six vs. nine elements. In both cases we need to count to fifteen. If we assume that people used the described strategy, we may then conclude that it is not the counting process that may explain the obtained difference between distances, but rather the comparison/integration process related to working memory resources. As we mentioned above, in this process the two cardinalities are being compared, and according to Dehaene (1999), the ANS mechanism is then automatically engaged. As a result, we observe the distance and size effects.

## General Discussion

We studied the cognitive mechanisms involved in the verification of proportional quantifiers. The first study showed that to verify proportional sentences substantially more time is needed than to only count and memorize relevant sets of objects. This was true when semantic condition was relatively simple in terms of the number of presented objects. However, along

with increasing total number of elements, we observed the speed accuracy trade off, which suggested that people took less time to process proportional judgments, in comparison with simpler conjunction structures, sacrificing their correctness. In the second investigation we strengthened the conclusion that proportional sentences require more cognitive resources than other sentences, because they were substantially associated with the cognitive control and memory storage, while their simpler counterparts were related only with the latter function. Finally, in the third study we explored the nature of the integration process specific for proportional quantifiers. We found that the numerical distance between groups of elements influenced the verification time; that is to say, the lower the distance, the slower the reaction. We hypothesized that in the comparison process well known phenomena related to the approximate number system may play a crucial role. Specifically, we suggested that small numerical distance between the two compared numbers requires more time (cf. Dehaene 1999).

The present studies have shown that the processing of proportional quantifiers is significantly more demanding than simple tracking, counting, and memorizing groups of objects. This suggests that the integration/comparison process may account for the uniqueness and difficulty of this type of quantifiers (McMillan et al. 2005; Szymanik and Zajenkowski 2010a). One may wonder whether the mechanisms potentially involved in the verification of proportional propositions are universal for complex verbal structures. Interestingly, many researchers have emphasized the role of WM for sentence comprehension. Since Miller and Chomsky's (1963) seminal study, it has been suggested that verbal memory is especially significant for comprehension of syntactically complex sentences. The authors observed that the embedded syntactic structures (e.g. "This land and these woods can be expected to rent itself and sell themselves, respectively"), though grammatically correct, for many people are difficult to understand. Such complex sentences require maintaining their parts in memory (e.g. noun phrases in the example given above), while trying to integrate them with other expressions (e.g. verbs). Other well-studied phenomena include grammatical dependencies such as the relations between words that cannot be adjacent. This is expressed for example in a relative clause of the form in (1).

(1) It was the writer that the professor wants to introduce to the editor.

To comprehend (1) and understand "Who is doing what to whom?" we should hold on to the noun "the writer" across intervening material, in particular nouns of similar type, such as "the professor", until it can be integrated with the object of the verb "to introduce". In a sentence such as (1), where the linear order of the arguments is in conflict with the structure of the sentence ("the professor" is linearly nearer to the verb "introduce" compared with "the writer"), the system is prone to error. The embedded sentences and object relative clauses, likewise proportional sentences, require holding in memory some parts of the sentence for further adequate integration with other expressions (e.g. Miller and Chomsky 1963; Gordon et al. 2002, 2004).

From an algorithmic perspective, the computational model of language production and comprehension, and the computational model of quantifier verification are essentially the same. The computational parallels between those tasks go beyond formal similarity. It seems that in both cognitive tasks—language comprehension and quantifier verification—we observe increase in WM activation when we consider the problem of higher-complexity. Context-free sentences place higher cognitive constraints on processing than regular constructions, and quantifiers corresponding to PDAs are more difficult to verify than quantifiers corresponding to finite-automata. This points toward an emerging view in cognitive science that the computational complexity of the cognitive task may correspond to its cognitive difficulty (Ristad 1993; Van Rooij 2008; Szymanik 2010).

On the other hand, we may try to identify general mechanism involved in the cognitive processing of complex sentences. In all the cases discussed above, namely embedded structures, object relative clauses, and proportional quantifiers, storage capacity as well as active integration of information in face of the interference is required. According to some researchers, dealing with interference is a part of working memory related to central executive or attentional control (Engle et al. 1999). The question arises whether this aspect is crucial for the processing of complex sentences, and if so, what is its nature. Research exploring central executive and cognitive control show that these concepts cover a wide range of functions. Miyake et al. (2000), for instance, identified three specific processes underlying central executive. First, the authors suggested that it inhibits dominant and irrelevant responses interfering with the contents of memory. Second, the central executive is responsible shifting between various tasks. Finally, the central executive monitors the information currently being held in memory and updates the contents which no-longer relevant. It remains an open question which of these functions are crucial for complex sentences processing.

**Acknowledgments** The work of the first author was supported by a Grant No. 2011/01/D/HS6/01920 funded by the National Science Centre in Poland. The second author would like to acknowledge a generous support of NWO Veni Grant 639.021.232.

## References

- Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). New York: Cambridge University Press.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1969). *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.
- Clark, R., & Grossman, M. (2007). Number sense and quantifier interpretation. *Topoi*, 26, 51–62.
- Dehaene, S. (1999). *The Number sense: How the mind creates mathematics*. Oxford: University Press.
- Ebbinghaus, H. D., Flum, J., & Thomas, W. (1996). *Mathematical logic. Undergraduate texts in mathematics*. New York: Springer.
- Engle, R., Kane, M., & Tuholski, S. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). London: Cambridge Press.
- Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory load interference in syntactic processing. *Psychological Science*, 13, 425–430.
- Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, 51, 97–114.
- Hackl, M. (2009). On the grammar and processing of proportional quantifiers: Most versus more than half. *Natural Language Semantics*, 17, 63–98.
- Halberda, J., Mazocco, M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, 455, 665–668.
- Hopcroft, J. E., Motwani, R., & Ullman, J. D. (2006). *Introduction to automata theory, languages, and computation*. Boston: Addison-Wesley Longman Publishing Co.
- Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122–149.
- Logie, R. H. (2011). The functional organisation and the capacity limits of working memory. *Current Directions in Psychological Science*, 20, 240–245.
- McMillan, C., Clark, R., Moore, P., Devita, C., & Grossman, M. (2005). Neural basis for generalized quantifiers comprehension. *Neuropsychologia*, 43, 1729–1737.
- McMillan, C., Clark, R., Moore, P., & Grossman, M. (2006). Quantifier comprehension in corticobasal degeneration. *Brain and Cognition*, 62, 250–260.
- Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In D. R. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. II). New York: Wiley.

- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.
- Pietroski, P., Lidz, J., Hunter, T., & Halberda, J. (2009). The meaning of ‘most’: Semantics, numerosity, and psychology. *Mind and Language*, *24*, 554–585.
- Redick, T., & Engle, R. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology*, *20*, 713–721.
- Ristad, E. S. (1993). *The Language complexity game. Artificial intelligence*. Cambridge, MA: The MIT Press.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652–654.
- Szymanik, J. (2007). A comment on a neuroimaging study of natural language quantifier comprehension. *Neuropsychologia*, *45*, 2158–2160.
- Szymanik, J. (2009). *Quantifiers in TIME and SPACE. Computational complexity of generalized quantifiers in natural language*. PhD thesis, University of Amsterdam.
- Szymanik, J. (2010). Computational complexity of polyadic lifts of generalized quantifiers in natural language. *Linguistics and Philosophy*, *33*, 215–250.
- Szymanik, J., & Zajenkowski, M. (2009). Improving methodology of quantifier comprehension experiments. *Neuropsychologia*, *47*, 2682–2683.
- Szymanik, J., & Zajenkowski, M. (2010a). Comprehension of simple quantifiers. Empirical evaluation of a computational model. *Cognitive Science*, *34*, 521–532.
- Szymanik, J., & Zajenkowski, M. (2010b). Quantifiers and working memory. *Lecture Notes in Artificial Intelligence*, *6042*, 456–464.
- Szymanik, J., & Zajenkowski, M. (2011). Contribution of working memory in the parity and proportional judgments. *Belgian Journal of Linguistics*, *25*, 189–206.
- Van Benthem, J. (1986). *Essays in logical semantics*. Dordrecht: Reidel.
- Van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science*, *32*, 939–984.
- Zajenkowski, M., Styła, R., & Szymanik, J. (2011). A computational approach to quantifiers as an explanation for some language impairments in schizophrenia. *Journal of Communication Disorders*, *44*, 595–600.