Contextualising generic and universal generalisations: quantifier domain restriction and the generic overgeneralisation effect

Dimitra Lazaridou-Chatzigoga[a,b,1] , Linnaea Stockall[c], Napoleon Katsos[a]

[a]Department of Theoretical and Applied Linguistics, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DA, United Kingdom
[b]Department of English and American Studies, Humboldt University of Berlin, Unter den Linden 6, 10099 Berlin, Germany
[c]Department of Linguistics, SLLF, Queen Mary, University of London, Mile End Road, London, E1 4NS, United Kingdom

Abstract

Generic generalisations (e.g. 'tigers have stripes', 'ducks lay eggs') refer to a characteristic property of a kind. Recently, the generics-as-default view has posited that we have a bias towards interpreting universally quantified statements as generic. Evidence offered for this view is the Generic Overgeneralisation (GOG) effect, which refers to the documented tendency of participants to misinterpret a quantificational statement like 'all ducks lay eggs' as if it were a generic and thus accept it as true, even though they know it is false. Across two experiments in English and Greek we systematically addressed the relevance of context and quantifier domain restriction for this kind of behaviour. Participants judged generic majority characteristic statements like 'tigers have stripes' or statements with universal quantifiers with different sensitivity to quantifier domain restriction preceded by one of three levels of context (neutral, contradictory and supportive). We found that context significantly affected the rates at which participants accepted universally quantified statements. Our results demonstrate that quantifier domain restriction is a viable alternative explanation for a significant proportion of the judgements of universally quantified statements that have been called GOG errors.

---

[1] Corresponding author at: Institut für Anglistik und Amerikanistik, Humboldt Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany.
E-mail address: dl518@cam.ac.uk (D. Lazaridou-Chatzigoga).

1. Introduction

Quantificational generalisations, as in (1)-(2), are expressed in quantitative, statistical terms, while generic generalisations, as in (3)-(4), make general claims about kinds of entities and refer to a property that is characteristic of the kind in question, but not necessarily statistically prevalent, as in (4):

(1)   Some lions live in cages.
(2)   All lions eat meat.
(3)   Lions roar.
(4)   Lions have manes.

Generic generalisations have long been studied in formal semantics, within which genericity is frequently viewed as a species of quantification involving a covert sentential operator called 'GEN' (Krifka, Pelletier, Carlson, ter Meulen, Chierchia & Link, 1995). However, how to characterise their semantic interpretation and how to model their truth conditions remain controversial topics (see Dahl, 1975; Carlson, 1977; and more recent discussion in Mari, Beyssade, & del Prete, 2013).

In recent years, a growing body of experimental and developmental psychological work on the topic proposes that genericity is categorically different from (and significantly simpler than) quantification (Leslie, 2007, 2008; Gelman, 2010). This latter hypothesis, called the *generics-as-default* view, treats generics as a cognitive default and argues that they have priority both in terms of ontogeny (children understand and produce generics before quantified statements; see e.g. Hollander, Gelman & Star, 2002; Gelman, 2010) and in terms of cognitive complexity (quantified statements are misunderstood or misrecalled more often than generics in experimental tasks, see e.g. Leslie, Khemlani, & Glucksberg, 2011; Leslie & Gelman, 2012). This view posits a 'generic bias', which is founded on a dual view of cognition that assumes a distinction between fast, automatic and effortless System 1 and slow, effortful, higher-level and rule-governed System 2 (Kahneman & Frederick, 2002). One piece of evidence for the existence of two systems is the fact that they can lead to conflicting judgments. Leslie (2007:395) cites Frederick's (2005) "cognitive reflection test", to illustrate the two systems: "A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?" Most people report an initial inclination to answer "10 cents". System 1 supplies this first fast, but erroneous response. The correct response of "5 cents"

requires algebraic reasoning, part of the slower System 2. Leslie (2007:397) proposes that generics are part of System 1 and quantifiers part of System 2:

> The evidence surveyed so far suggests that System 1 - the more primitive system - is not particularly sensitive to information about how much or how many. I suggest that generics are judgments issued by System 1. They are thus non-quantificational; they do not depend on considerations of quantity, or any such information easily captured by set-theory. They are, however, automatic, effortless, and cognitively basic. Quantifiers, in contrast, express judgments issued by System 2, the rule-governed, extension-sensitive, higher-level system. Quantifiers do depend on considerations such as how much and how many. They are thus easily describable in the terms of set-theory.

Inspired by other cases of System 1 overuse like the one mentioned above (Frederick, 2005), Leslie (2007) anticipates similar errors in the interpretation of generic (System 1) and quantificational (System 2) statements, which would take the shape of a 'generic bias'. According to the generics-as-default proponents, evidence for the suggested 'generic bias' could come from different sources. Two main examples of overuse of System 1 that leads to overgeneralisation of quantifiers as generics have been described in the literature: (a) 3-year-old children treat quantified statements with 'some' and 'all' as if they were generic more often than the reverse (Hollander et al., 2002) and (b) adults misunderstand universal quantificational statements as generics and show a tendency to endorse false universal generalisations when the corresponding generic is true (e.g. they accept 'all ducks lay eggs' as true despite knowing that male ducks do not do so; Leslie et al., 2011). This led Leslie et al. (2011) to proclaim the Generic OverGeneralisation effect ('GOG' henceforth) as one of the main pieces of evidence in support of the generics-as-default view. The GOG effect is defined as "the tendency to overgeneralise from the truth of a generic to the truth of the corresponding universal statement" (Leslie et al. 2011, p. 17). This tendency is attributed to generics' default nature (see Prasada, Khemlani, Leslie, & Glucksberg, 2013, for more discussion of this conceptually based approach to generics).

Furthermore, according to this view, the fact that no language has a dedicated overt 'GEN' operator (Dahl, 1995) does not come as a surprise: given that generics are the most primitive default generalisations, children do not need to learn anything special in order to acquire them. Generics come essentially for free. In contrast, effortful, non-default quantificational generalisations require overt linguistic expression. However, while assigning

generics to a more basic, unmarked System 1, mode of thinking may sound intuitive at some level, it rests on a vague and undefined notion of markedness. Leslie (2008) cites a pattern discussed in Chomsky (2000), but she never spells out precisely what formalisation of markedness she relies on. Intuitively, it seems that what is at stake is surface level overt realisation (the third notion of markedness in Haspelmath, 2006).

In previous work (Lazaridou-Chatzigoga, Stockall, & Katsos, 2017) we considered the conceptual arguments for the generics-as-default hypothesis and the status of the GOG effect as a processing error. We concluded that evidence that the GOG responses are erroneous judgments is less compelling than it might at first appear.

The goal of this paper is to provide empirical evidence that the pragmatic phenomenon of quantifier domain restriction (QDR henceforth) is an alternative viable explanation for a substantial part of the attested behaviour, that is, of the fact that adults accept statements like 'all tigers have stripes' and 'all ducks lay eggs' as true. Our rationale is that in order to evaluate the theoretical and psychological validity of a new mechanism for the interpretation of quantified statements, such as the GOG effect, one needs to carefully test and consider the already available alternative explanations, which rely on independently established factors. We will show that context affects the rates at which participants accept universally quantified statements as true in a way that relies on QDR and on the differing sensitivity different universal quantifiers show to it. Our results show that the attested behaviour can be largely explained by QDR, a phenomenon about which we have abundant independent evidence. The size of the QDR effect in our experiments shows that it is a major contributing factor to the attested behaviour. If there is a GOG effect then, as argued in much recent literature, it is smaller than was originally proposed. However, QDR does not explain all of the prima facie 'mistaken' acceptances of universal statements and therefore there is still some residual behaviour to be accounted for, so we also make a broader point: In order to properly evaluate the contribution of the GOG effect, we need to engage in quantitative linguistically-informed research that takes into account all the other known factors that lead to GOG-like behaviour, since QDR is only one of them.

We ran the same experiment in English and Greek and confirmed that our context manipulation worked in two different languages. By doing so, we also addressed the unmarkedness claim for generics, as the generic statements mostly studied thus far were bare plural generics in English, which lend themselves easily to the idea of unmarkedness, as they are characterised by the absence of a determiner. In Greek, definite plurals are used instead of bare plurals making generics look less 'unmarked' and, moreover, the configuration of

nominals involved in genericity, quantification and definiteness differs from English. Our results point further to the significance of doing experimental work on genericity and quantification cross-linguistically.

2. The GOG effect: background and alternative explanations

The generics-as-default view argues that people have the tendency to interpret quantified statements as if they were generic, e.g. evaluating a statement quantified with 'all' or 'some' as though it were a generic. Support for this view is based on data from universal and existential statements from English, Mandarin and Quechua-speaking children (see e.g. Hollander et al., 2002; Tardif, Gelman, Fu, & Zhu, 2012; Mannheim, Gelman, Escalante, Huayhua, & Puma, 2011 for each language respectively)[2] and on data from universal statements from English-speaking adults, under certain circumstances (see Khemlani, Leslie, Glucksberg, & Rubio-Fernandez, 2007; Leslie et al., 2011; Meyer, Gelman, & Stilwell, 2011).

The first detailed investigation of the scope of the GOG effect in adults is found in Leslie et al. (2011). In their experiment 1, participants performed a truth-value judgement task on sentences that were presented in generic, universal ('all'), or existential ('some') form. The statements involved different kinds of properties: quasi-definitional ('triangles have three sides'), majority characteristic ('tigers have stripes'), minority characteristic ('ducks lay eggs'), majority non-characteristic ('cars have radios'), striking ('pit bulls maul children'), and false generalisations ('Canadians are right-handed'). Leslie et al. report that adults sometimes judged universal statements as true, despite knowing that they were truth-conditionally false. For example, participants judged a statement like 'all tigers have stripes' as true, even though it is false given that there are albino tigers, and they accepted 'all ducks lay eggs', even though only sexually mature female ducks have this capacity. Leslie et al. claim that the participants made this 'error' because they relied on the corresponding generic statements, which are true ('tigers have stripes', 'ducks lay eggs'). The authors find that the GOG effect is restricted to *characteristic* properties and that it occurs in more than half the trials: 78% for majority characteristic and 51% for minority characteristic statements.

Leslie et al. entertained three alternative explanations, which they argue are ruled out with subsequent experiments: a) ignorance of the relevant facts, b) subkind interpretation,

---

[2] See Lazaridou-Chatzigoga, Katsos, & Stockall (2015) for a critical review of the child language evidence.

and c) QDR. Thus, Leslie et al. (2011) claimed to have evidence for a generic bias, according to which people sometimes treat universally quantified statements as if they were generic.

In previous work (Lazaridou-Chatzigoga, Stockall, & Katsos, 2017) we provided conceptual arguments against the rejection of the above alternative explanations, and also considered a fourth explanation, the atypical behaviour of 'all'. Here is a summary (alongside our criticism):

a. *ignorance of the relevant facts*: participants might simply be ignorant of the relevant facts that mainly involve primary or secondary sexual characteristics relevant to mating and reproduction. For instance, for 'all ducks lay eggs' participants might not know that male ducks do not lay eggs. Leslie et al. addressed this through a knowledge task (experiments 3a,b) and they were able to confirm that participants knew the facts. Interestingly, though, acceptance of statements like 'all ducks lay eggs' fell to 33% when the knowledge task was performed before judging the critical statements (compared with 51% in the original experiment). While Leslie et al. are correct to point out that these statements were still accepted as true at a higher rate than they ought to be, we take an 18% reduction as indicative that alternatives to the GOG explanation are worth pursuing.[3]

b. *subkind (taxonomic) interpretation*: participants might be interpreting 'all ducks lay eggs' as 'all kinds/types of ducks lay eggs', including the Mallard, the Eider, the Goldeneye, etc. Under such a subkind interpretation, the correct response is to accept them, thus there is no need to look for explanations of that behaviour. Leslie et al. rejected this explanation based on their experiment 2b, where participants were asked to provide a paraphrase of each statement trying to keep the paraphrase as close to the meaning of the original statement as possible. Participants used subtyping language in only 1% in their paraphrasing. The fact, though, that participants were not *conscious* of the fact that they might have interpreted the statement as involving subkinds is non-conclusive, as the distinction between implicit knowledge and explicit knowledge is a fundamental one in cognitive science. For instance, Dienes & Perner (2002) discuss various cases of participants employing rules in categorisation and judgment tasks, which they are unable to make explicit when asked. Further research actually

---

[3] For some puzzling effects related to statements like 'tigers have stripes' when participants did the knowledge task before evaluating the statements, see our discussion Lazaridou-Chatzigoga, Stockall, & Katsos (2017).

manipulating the availability of a subkind interpretation is required before rejecting this explanation.

c. *quantifier domain restriction*: participants might be interpreting a statement like 'all ducks lay eggs' as applying only to a relevant subset of ducks, namely the mature fertile female ducks. This is based on the assumption that quantified statements are interpreted within a context, which may restrict the scope of the quantifier (as per Stanley & Szabó, 2000; Stanley, 2002). This alternative explanation was addressed in Leslie et al.'s experiment 2a, where they provided the participants with a context which supplied population information (e.g. "Suppose the following is true: there are 431 million ducks in the world. Do you agree with the following: all ducks lay eggs"). The GOG effect dropped by 18% for majority characteristic and by 21% for minority characteristic statements with respect to their experiment 1, but it still occurred on a substantial portion of trials for statements with 'all' (60% for majority and 30% for minority characteristic). QDR was also investigated through their experiment 2b, which included a paraphrase task. Of the paraphrases provided by the participants only 1.6% included a subset interpretation (e.g. "female ducks lay eggs"). The same criticism outlined in (b) above applies here with respect to the distinction between implicit and explicit knowledge. Thus, Leslie et al. discarded QDR as a major alternative explanation for the GOG effect. For more discussion of QDR see section 2.1 below.

d. although not an explanation considered by Leslie et al. (2011), a fourth explanation could be the *atypical behaviour of* 'all': 'all' cannot be treated as a representative universal quantifier. It has been argued that a) 'all' participates in fallacious reasoning (Jönsson & Hampton, 2006), b) 'all' is prone to hyperbolic/loose use with a meaning similar to 'very many' or 'almost all' (see Claridge, 2011, for work on corpus linguistics), c) 'all' is ambiguous between distributive and collective interpretation (Beghelli & Stowell, 1997)[4]. Thus, we argue that other universal quantifiers in English and other languages can offer a better test of the generics-as-default hypothesis.

---

[4] A further potential licensor of a 'true' response is the possibility of an association between the concept TIGER and the concept STRIPE, that is, 'all tigers have stripes' might be judged as true, because the concept TIGER is strongly associated with the concept STRIPE. For instance, in a structural model of semantic memory (Collins and Quillian, 1969) the concept TIGER would be linked to the concept STRIPE or in a model based on semantic features the feature STRIPE would be part of the meaning of the word TIGER (Smith, Shoben, and Rips, 1974). See Hampton (2015) for further details on these models as well as discussion of their shortcomings.

In sum, Leslie et al.'s (2011) dismissal of the above explanations is not as convincing as they argue. We proposed that ignorance, subkind interpretation, atypical behaviour of 'all' and QDR might all play a role in explaining the attested behaviour by adults (ignorance and QDR were responsible for a 18/21% reduction in 'true' responses respectively and subkind interpretation and QDR were found in paraphrases at 1/1.6% rates respectively in Leslie et al.'s own experiments). If the effects of each of these factors are independent and cumulative, the residue of real errors in Leslie et al. (2011)'s experiment 1 left to explain as GOG is 40% or less out of the 78% for majority characteristic statements, and as low as about 10% out of the 51% for the minority characteristic statements. We also argued that even the name of the GOG effect might be misleading. The effect mainly tries to capture the behaviour observed with 'all' only when the property in question is characteristic, which supposedly receives a generic interpretation as a result of an overgeneralisation bias. Thus, perhaps a better name for that effect would be 'Quantifier Reinterpretation' effect, because this term would direct the focus where we believe it belongs: on the interpretation of 'all', or more generally of quantifiers, rather than the interpretation of generic statements.

In this paper, we look for empirical evidence to test these conceptual arguments, focusing on QDR and its relevance for the attested behaviour.

3. Interpreting universally quantified generalisations

3.1 The behaviour attributed to the GOG effect within the generics-as-default view

As mentioned above, Leslie et al. (2011) addressed QDR as an alternative explanation for why participants accept statements like 'all tigers have stripes' or 'all ducks lay eggs'. The contexts used by Leslie et al. (2011) relied on population information (e.g. "Suppose the following is true: there are 431 million ducks in the world. Do you agree with the following: all ducks lay eggs"). This information was supposed to prime quantification over every individual duck in the world, and thereby to make it difficult/impossible to interpret 'all' as restricted to only the ducks that are presupposed by 'lay eggs'. If acceptance of 'all ducks lay eggs' without any context was driven by QDR, Leslie et al. predicted that it would disappear in the context of population information. These contexts though only had a moderate effect on participant behaviour. While the GOG effect dropped with respect to the rates they obtained in an experiment without any context, there was still a high acceptance of such statements.

Leslie et al. (2011) argue that "the effect cannot be entirely explained this way [i.e. via QDR], since the effect remained in a context that encouraged generalization across each and every individual in the world". While they only find the effect with characteristic statements and with the universal quantifier 'all', they often discuss their results as if they applied to all universal statements (e.g. in Leslie et al. 2011, p. 18: "the generics-as-default hypothesis predicts that adults will tend to incorrectly endorse false universal statements if the corresponding generic is true"; "the tendency to substitute a judgment of the generic for the universal will be generally more successful when the property in question is a characteristic one").

Leslie et al. (2011) thus discarded QDR as a contributing factor to the attested behaviour. On the basis of the additional experiments they conducted to address alternative explanations, they concluded that this behaviour can only plausibly be explained if we attribute it to Leslie's (2008) and Gelman's (2010) hypothesis that generics express primitive, default generalisations. Thus, it is generics' default nature that leads participants to misinterpret universally quantified statements as if they were generic. Treating this behaviour as an error is an integral part of this approach, as this lends support to the existence of a strong generic bias that is expected given that generics involve more basic (System 1) generalisations, while universally quantified generalisations concern non-default, more effortful ones.

If one were to design a study addressing the relevance of context via manipulating different levels of it, on our understanding of the generics-as-default view, proponents of that view would expect that different kinds of context should give rise to similar levels of GOG responses, for all universal quantifiers. The 'more marked' system 2 quantified statements should be 'erroneously' interpreted as generic statements at similar rates to previous experiments. Proponents of the generics-as-default view would expect only very moderate differences between different kinds of context, or between different quantifiers, since the generics-as-default view is that participants fail to interpret the quantifiers as universal quantifiers, and that they erroneously accept the generic statement counterparts instead. In sum, the generics-as-default view proposes that the attested behaviour is due to an underlying generic bias that manifests as a failure to correctly process the quantifier and not to any specific properties of the quantifier itself. Thus, in principle it should make no difference what the quantifier is, and this view is silent about how to explain any potential differences one might observe between different universal quantifiers as a function of context.

3.2 Quantifier domain restriction as an alternative explanation

We argue that QDR needs to be investigated as a viable alternative explanation for much of the data attributed to the GOG effect, building on a design used by Lazaridou-Chatzigoga & Stockall (2013). We focus on QDR among the other alternative explanations given that it is a pervasive phenomenon affecting quantifiers and their interpretation within a context, and is routinely invoked in quantification (von Fintel, 1994). According to QDR, the domain of a quantifier can be restricted: in a discourse like *'There was rhubarb pie for dessert. Everyone developed a rash'* (example modified from von Fintel, 1994, p. 33), 'everyone' does not quantify over all the individuals in the world, but rather over the contextually restricted set of individuals who ate the rhubarb pie. Furthermore, listeners are known to be charitable (Grice, 1975). Thus, in a conversation one assumes that speakers take the most sensible positions and make the most plausible assertions. Under this view, interpreting 'everyone' as quantifying over all the individuals in the world seems a rather unlikely intended interpretation and moreover one that is not charitable to the speaker because it renders her utterance false, whereas interpreting 'everyone' with respect to the available set of individuals is not only plausible but also charitable to the speaker. How we encode QDR in the grammar is currently under debate and opinions vary as to whether QDR is part of the syntax/semantics (Stanley & Szabó, 2000; Stanley, 2002, von Fintel, 1994; Martí, 2003; Giannakidou, 2004), or of the pragmatics (see e.g. Recanati, 1996) - this is not the place to decide upon this issue (see Kratzer, 2004). Irrespective of whether one represents QDR as a syntactic, semantic or pragmatic phenomenon, there is an abundance of independent motivation for its existence.

To come back to the statements relevant for our studies, if QDR was relevant when participants interpreted statements like 'all ducks lay eggs' or 'all tigers have stripes' it would mean that participants would interpret these statements as applying only to a relevant subset of the kind, that is, the mature fertile female ducks in the case of 'all ducks lay eggs' and the normal tigers in the case of 'all tigers have stripes'. If participants are indeed restricting the domain to these relevant sets, then their behaviour is not erroneous, as they are indeed expected to accept these statements under a QDR interpretation.

We hypothesised that if we could show that the rate at which participants 'erroneously' accept universally quantified statements (the behaviour labelled the GOG effect) can be altered by carefully manipulating different levels of contextual information preceding the critical utterance, we would have evidence that the observed tendency to accept universally quantified statements as true can be largely explained through QDR, an

independently motivated mechanism, and that resorting to an explanation such as the GOG would be less appealing.

The contexts used by Leslie et al. (2011) relied on population information (e.g. "Suppose the following is true: there are 431 million ducks in the world. Do you agree with the following: all ducks lay eggs") and only had a moderate effect on participant behaviour. We argue that these contexts did not succeed in making salient the potential restricted domain against which participants might be interpreting the statements, that is, they did not make the set of female ducks salient. Instead, they provided participants with an estimate about the total number of ducks in the world. Our contexts would crucially include information that explicitly related to the potential restricted domains. We provide an overview of our studies in the next section.

4. Overview of the studies

In the present studies we investigated QDR as a viable alternative explanation for much of the data attributed to the GOG effect, building on a design used by Lazaridou-Chatzigoga & Stockall (2013). Because of the design we adopted, we focused only on majority characteristic statements ('tigers have stripes') leaving minority characteristic statements ('ducks lay eggs') for future investigation. The majority characteristic items evoked a much larger 'GOG' effect in the original study, and they moreover constitute the more general case. Minority characteristic statements seem to exclusively involve primary or secondary sexual characteristics of animal kinds. Likely due to this restriction, of the twelve minority characteristic items in the original study by Leslie et al. (2011), the same predicate was repeated either two ('deer have antlers', 'moose have antlers') or three times ('ducks lay eggs', 'insects lay eggs', 'snakes lay eggs'). Therefore, if we excluded second and third occurrences of a predicate as well as one item that was removed post-hoc from their analyses ('cardinals are red'), we would have been left with only 8 items in our item-base. Our own attempts to construct additional, non-repeated items proved difficult because of the need to resort to more specialised vocabulary such as 'udders' or 'foals', or to rely on knowledge unlikely to be broadly shared[5]. Thus, we decided to leave minority characteristic items out of this study.

---

[5] For example, in pre-testing a possible candidate set of items, we discovered considerable variation in the knowledge of London-based undergraduates about whether pigs suckle their young or cardinals are red, etc.

In our design, we manipulated context and investigated the relevance of QDR for the attested behaviour. We reasoned that when people accept a statement like 'all tigers have stripes' they interpret it as a claim only about the relevant restricted set of normal instances of the kind, which are characterised by having stripes, as the counterexamples may not be spontaneously accessible. We decided to use three levels of contexts and varied the context preceding the critical utterance as follows: a) neutral, where the information in the context does not interact with the truth value of the critical statement, b) contradictory, where exceptions which should rule out a universally quantified statement are made salient, and c) supportive, where the generality of the critical property is made salient through a paraphrase. Examples of each context level are given below:

(5)   a. **neutral**: Linton Zoo is home to three tigers, Tibor, Baginda and Kaytlin, whose playful games visitors love to watch and photograph.
b. **contradictory**: Linton Zoo is home to three tigers, Tibor, Baginda and Kaytlin, whose fur is all white due to a recessive gene that controls coat colour.
c. **supportive**: Linton Zoo is home to three tigers, Tibor, Baginda and Kaytlin, whose black and orange coats visitors love to photograph.

The contradictory and supportive contexts would make the relevant domain for QDR salient, while the neutral context would serve as a baseline measure. The contradictory and supportive contexts turned the implicit restriction to 'all normal' individuals into an explicit one by either highlighting some abnormal individuals (contradictory) or by using a paraphrase that suggested that the relevant individuals had the relevant property, i.e. they were normal individuals (supportive).

As for generics, standardly they are argued to resemble adverbial quantification and to differ from nominal quantifiers in that they resist contextual narrowing to a salient set of entities. The following examples (Krifka, 1987, p.7) show that while the nominal argument of 'every' ('lion' in (6)), is subject to QDR and can refer to the set of lions introduced in the context, this is not a possible interpretation for the indefinite singular 'a lion' in (6), which expresses a property of lions in general. Thus, the truth conditions of generics seem to remain stable across different contexts, while the truth conditions of quantified statements vary depending on the context of utterance.

(6)   a. (Out of the blue): Every lion has a mane. (non-restricted)

b. There are lions and tigers in the cage. Every lion has a mane.
   (restricted or non-restricted)

c. There are lions and tigers in the cage. A lion has a mane.
   (non-restricted only)

d. There are lions and tigers in the cage. A lion always has a mane.
   (non-restricted only)

Following the standard view of generics that considers them to be immune to context and given that the generic statements were chosen to be 'true', participants were expected to accept generics as true at equally high rates in all contexts – even in the contradictory context given generics' tolerance of exceptions. Intuitively though, the saliency of the exceptions might give rise to lower acceptance rates in the contradictory context compared to the other conditions, a result that would align with other proposals in the literature that, contrary to the standard view, posit context sensitivity for generics (Pelletier and Asher, 1997; Cohen 2001; Nickel, 2008; see Sterken, 2015 for an overview). This sensitivity is based on examples like 'dobermans have floppy ears', which is true in the context of evolutionary biology, but false in the context of dog breeding (Nickel, 2008). According to these views we might expect some context sensitivity even for generics.

Note though that the contrast in (6) does not obtain for all universal quantifiers equally. Universal quantifiers are not equally subject to QDR in all contexts. 'All', 'all the', 'every' and 'each' are all universal quantifiers, but they differ in terms of distributivity (Vendler, 1962; Gil, 1995), and on whether they allow, may resist or require QDR. Using different types of universal quantifiers is essential to test the scope of a QDR-based explanation of the GOG effect, as universal quantifiers show different sensitivity to QDR: QDR is less likely if the universal quantifier does not require/may resist linking with a set under discussion (Partee, 1995; Matthewson, 2001), as is the case with 'all', compared to 'all the' and 'each', which have to be interpreted as D(iscourse)-linked (Pesetsky, 1987). There seems to be essentially an ordering of preference with respect to QDR: 'all the' and 'each' require it, while 'all' allows or may resist it, as the following examples illustrate:

(7)   Context: There are three cats in the garden, which happen to be black.
       a. All cats are black.
       b. All the cats are black.
       c. Each cat is black.

In the above context, (a) would most probably be assigned a non-restricted interpretation and would be treated as false. It is possible though for (a) to be treated as true, if one interprets (a) as referring to 'all cats in the garden'. According to Matthewson (2011), 'all' is not susceptible to QDR and such an interpretation is expected to be infelicitous. We believe though that this is an empirical question that needs to be tested. In the above context, both (b) and (c) should be treated as true, given that 'all the' and 'each' necessarily have to be interpreted with respect to the context of utterance, which happens in this case to render them true.

These differences among universal quantifiers motivate the use of several of them in an experiment addressing QDR. Using different universal quantifiers would also address the possibility that the atypical nature of 'all' plays a part in generating the attested GOG behaviour.[6] Thus, in addition to manipulating context, a compelling test of the QDR view also requires testing whether the GOG effect is observed only with 'all'. As we discussed above, there are reasons to believe that 'all' is not a representative universal quantifier. Only Lazaridou-Chatzigoga & Stockall (2013) have tested universal generalisations other than 'all', obtaining differences between 'all', 'every' and 'all the' that supported further investigation of different types of universal quantification.

In the neutral condition, the predictions of the QDR account critically diverge from the generics-as-default view, as we expect that participants' responses are not biased towards a generic interpretation, but are rather dependent on the sensitivity of the quantifier to QDR and on the available non-QDR alternatives to accepting these statements. The neutral context makes some individuals of a kind salient, i.e. a set of tigers, without providing any information relevant to the property in question. The neutral context will provide a baseline measure in order to observe any differences between quantifier types, while keeping the information supplied by QDR constant. To judge whether 'all the' or 'each' statements are true in the neutral context, participants are required to access their own knowledge about the kind under discussion. We predict mostly 'true' responses, that is, that participants will accept the statements, given that the context does not provide any reason to doubt that the

---

[6] It would be also worth testing whether it is only D-quantifiers (that is, Determiner-quantifiers like 'all' and 'every'), that show this behaviour or whether A-quantifiers (that is, Adverbs, Auxiliaries, Affixes, etc. like 'always', 'usually', 'must' etc.) would be also affected. In principle, if the attested behaviour is due to a cognitive bias, we should expect to find it with all kinds of quantifiers. The relevance of QDR as a major alternative explanation for the attested behaviour could be tested more broadly if quantificational adverbs like 'always', 'usually' and 'typically' were tested as well. We leave that investigation for future research.

normal situation holds. Participants should rely on their previous experience of the world in order to give a judgment. Statements with 'all' in the neutral context could evoke acceptance of the statement for any of the reasons discussed above (ignorance, subkind interpretation, loose/hyperbolic interpretation). The availability of a wider range of possible interpretations for 'all' might give rise to higher acceptances compared to 'all the' and 'each'.

In the contradictory and supportive contexts, the QDR view predicts variation in acceptances: across the board, the contradictory context should decrease acceptance rates and the supportive context should increase them. In the contradictory context, the QDR view predicts low acceptance rates for 'all the' and 'each' statements, as participants are required to interpret the quantifiers as referring for instance to the stripeless tigers just introduced in the context. On the other hand, given that 'all' may resist a restricted interpretation, participants are not expected to adopt a QDR interpretation. Thus, 'all' statements could still trigger relatively high acceptances, as participants could still alternatively generate any of the other options (namely, the subkind interpretation or a hyperbolic/loose interpretation), which would license acceptance of the statement. They could not, however, be ignorant of the existence of stripeless tigers, so we would still predict that acceptances for 'all' would be lower in the contradictory context than in the neutral or supportive contexts.

In the supportive context, the QDR view predicts that, given that the QDR interpretation is easily available, and charitable, and because the non-QDR alternatives also license a 'true' response (subkind interpretation, loose/hyperbolic interpretation), acceptance rates for 'all' in the supportive context should be higher than in the neutral context. For 'all the' and 'each', explicitly predicating variegated coloration of the tigers under discussion means that the only licit judgement to the statements would be 'true', thus acceptances are expected to be higher than in the neutral context.

In our studies, we addressed generics cross-linguistically by running the same study in two languages, English and Greek. This enabled us to see whether the context manipulation in our design would work in a language different from English and furthermore addressed the following two issues: (a) the unmarkedness claim with respect to generics, and (b) the fact that Greek has a different repertoire of articles and quantifiers, i.e., the configuration of nominals involved in genericity, quantification and definiteness differs from English (Chierchia, 1998; Longobardi, 2001).

Early articulations of the generics-as-default hypothesis (Leslie 2007, 2008) appeal to a notion of markedness they attribute to Chomsky (2000). Leslie (2008, p. 24) discusses

Chomsky's example of 'John climbed the mountain' adopting the suggestion that "this is understood as meaning that John climbed *up* the mountain; to obtain the interpretation that John climbed *down* the mountain, we must explicitly use the preposition 'down'. The unmarked case 'climbed the mountain' is never interpreted as *climbed down the mountain*." The proposal is that we conceive of *climbing* as *climbing up* in the default, unmarked case, and have to use a more marked form to express the opposite.

Leslie then argues that by analogy "in the case of generalizations, the unmarked generic invokes the cognitive system's default mode of generalizing. To invoke a nondefault generalization, an explicit and marked quantifier must be used. The connection, then, between generic generalizations and unmarked surface forms is not an accidental one, but rather reflects a deep fact about human cognition. It is no coincidence that no human language has a dedicated, articulated generic operator". (Leslie 2008: 24)

This claim seems to rest on the observation that in English, generic generalisations are often expressed with statements that involve a bare plural subject and a simple present predicate and are thus shorter and morphologically simpler than quantified statements.

Except for the Chomsky citation, no definition of markedness is given. Intuitively, it seems that what is at stake is what Haspelmath (2006) discusses as formal markedness, i.e. surface level overt realisation of the type "in English, the past tense is marked (by -ed) and the present tense is unmarked". This seems to work well for English bare plural generics compared to quantifiers like 'all' (compare 'birds fly' to 'all birds fly'). However, it is not clear how it works for the other types of generics in English (indefinite singular, e.g. 'a cat has whiskers' and definite singular, e.g. 'the cat has whiskers') compared to quantifiers ('some cats have whiskers').

The claims that "there is no language that has a 'generic' article, that is, an article which is used exclusively with generic NPs" (Dahl, 1995, p. 425) and that "generics have the tendency to employ the least marked tense-aspect choice in the language" (Dahl, 1995, p. 415) have proven critical for theories of generics despite the fact that the typology of genericity has rarely been tackled systematically (see though Dahl, 1985; Behrens, 2000, 2005; Dayal, 2004). More importantly, in the generics-as-default view these two claims are erroneously confounded into a single claim that generics assume the most unmarked form in the *nominal* domain (Leslie, 2007). This proposal seemingly offers a nice solution to how English learning children can master the distinction between generic and universal or specific statements at an early age (Gelman, 2010, see Lazaridou-Chatzigoga, Katsos, & Stockall, 2015, for discussion). However, while intuitive, this appeal to markedness fails to extend to a

wide range of languages, many of which do *not* use a simpler form for expressing generic generalisations.

Without a clear and robust definition of markedness and a systematic examination of the form of generics alongside quantification and definiteness cross-linguistically, the appeal to unmarkedness and hence to defaultness for generics is impossible to evaluate. Focusing on languages that employ the presence/absence of a determiner in combination with tense/aspect features to express genericity, we observe that Greek differs from English in using a definite plural as the most frequent form for generics (see e.g. Marmaridou-Protopapa, 1984; Roussou & Tsimpli, 1994; Giannakidou & Stavrou, 1999, Lazaridou-Chatzigoga, 2009; Giannakidou, 2012), while bare plural generics in the preverbal position are generally disallowed:

(8) I   tighris ehun   righes.
     the.PL tigers   have.PL   stripes
     'Tigers have stripes.'

(9) *Tighris   ehun   righes.
      tigers   have.PL   stripes

Superficially, we already observe that the unmarkedness claim for generics does not hold cross-linguistically if we take it to be formal markedness in the sense of Haspelmath (2006). If we compare *i tighris ehun righes* ('tigers have stripes') to *kathe tighri ehi righes* ('every tiger has stripes') or to *merikes tighris ehun righes* ('some tigers have stripes') we observe that the generic statement is not less marked than a quantified statement. Not only are generics more 'marked' in Greek than in English (and in this respect, Greek is similar to Romance languages like French or Spanish, see e.g. Vergnaud & Zubizarreta, 1992; Longobardi, 2001; Chierchia, 1998; Dayal, 2004; Ionin and Montrul, 2010), but also *i tighris ehun righes* is potentially ambiguous between a generic and a definite (anaphoric) interpretation. Given that it can also mean 'the (contextually salient and previously mentioned) tigers have stripes', the generic form is not superficially differentiated from the anaphoric form.

Of course, one could also argue for another definition of markedness, according to which something is unmarked if it lacks a dedicated marker in a language. In this sense, Greek generics might still be argued to be unmarked as there is no dedicated marker of genericity in this language. Nevertheless, as we pointed out, without a clear and principled definition of (un)markedness, the appeal to unmarkedness and hence to defaultness for generics is

impossible to evaluate, and there is a danger of selecting whatever sense of markedness happens to fit a given dataset. Our Greek experiment (Experiment 2) offers a comparison between the processing of generic and universally quantified statements in a language where generalisations do not differ in formal markedness, thereby providing some initial data to address the issue.

To recapitulate, (a) the contextual manipulations used in our experiment were expected to make the implicit domain restriction explicit and salient to the participants, (b) this manipulation was expected to influence truth-value judgements of universal generalisations by showing a decrease in acceptance rates in the contradictory condition and an increase in acceptance rates in the supportive condition, and (c) the context effects were expected to be different for different determiners, depending on their semantics and their sensitivity to QDR. We also checked whether our context manipulation would work in a language other than English by running the same study in Greek. By doing so, we addressed one interpretation of the unmarkedness claim for generics and we highlighted the importance of doing experimental work on genericity and quantification cross-linguistically.

5. Experiment 1: Generalisations in English

5.1 Method

5.1.1 Participants and procedure

120 volunteers (49 male, 70 female, 1 other; aged 19-67; mean age 37.28; SD 13.06) participated in the experiment over the Internet. Participants were recruited through Amazon's MTurk system for human interface tasks. All spoke English as their first language and lived in the United States. All participants provided their informed consent and received payment. The study met the guidelines for ethical research with human participants of the Department of Theoretical and Applied Linguistics, University of Cambridge.
The study was presented in the online platform Qualtrics. Each trial consisted of three displays: (1) participants read a background context, (2) they read a statement, and (3) they were asked to judge whether they agreed with the statement they just read. Their response was recorded by selecting keyboard keys ('A' for yes; 'K' for no).

5.2 Materials and design

Participants judged 84 statements, including 48 fillers presented in a randomised order. The 12 test items consisted of majority characteristic statements like 'tigers have stripes' and 'horses have four legs'. We included 24 control items, 12 definitional statements like 'ants are insects' and 12 false generalisations like 'books are paperbacks' to obtain baseline measures and to (semi)-counterbalance the percentage of expected 'true'/ 'false' responses. All the contexts and items were normed beforehand by English native speakers, who did not take part in the experiment. Norming tested whether there was any unknown vocabulary and whether the context-statement pairings sounded natural, that is, if the manipulations worked as intended. Debriefing after norming clarified any additional comments. Some minor adjustments were made afterwards before the items were finalised for the final experiment. Most experimental items were a subset of the items used by Leslie et al. (2011). The two conditions we manipulated for the majority characteristic items were:

a. determiner type: bare plural generic/ 'all'/ 'all the'/ 'each'
b. context type: neutral/ contradictory/ supportive.

We varied the context preceding the critical utterance as follows: a) neutral, where the information in the context does not interact with the truth value of the critical statement, b) contradictory, where exceptions which should rule out a universally quantified statement are made salient, and c) supportive, where the generality of the critical property is made salient through a paraphrase. The statements were in one of the four determiner forms and were preceded by one of three levels of context (neutral/contradictory/ supportive), examples of which can be found in (5) above. We deliberately never used the same words to describe the relevant characteristic in our contexts and in the statement we asked participants to evaluate. For instance, our participants would have to infer from the claim in the contradictory context that the specific tigers' 'fur is all white' that they 'do not have stripes' or in the supportive context that the 'black and orange coats' is a paraphrase of 'have stripes'. We were keen to avoid priming the 'tiger/stripe' lexical association by repetition.

Given the 4 determiners (generic/ 'all'/ 'all the'/ 'each') we created 4 lists with 3 sublists each that varied with respect to the pairing of the items with context type, which gave us 12

lists in total. 10 participants were assigned randomly to each of the 12 lists. In (11) we see a sample of a trial of a statement with 'all' after a neutral context:

(10)

DISPLAY 1:
Background:

Linton Zoo is home to three tigers, Tibor, Baginda and Kaytlin, whose playful games visitors love to watch and photograph.

DISPLAY 2:
Statement: All tigers have stripes.

DISPLAY 3:
Do you agree with the statement?
o Yes (A)   o No (K)

The definitional and false generalisations were in the generic form in all lists. Fillers served to ensure the percentage of expected 'true'/ 'false' responses was similar. The definitional and false generalisations, as well as the fillers, were preceded by a context that did not vary across conditions. All materials can be found in Appendix A in the Supplementary Material.

5.3 Results and discussion

The final analysis included 116 participants. Four participants were excluded as they responded correctly to fewer than 10 out of the 12 definitional statements.

Table 1 summarises the proportion of 'yes' responses to the TVJ question for the test items (majority characteristic statements) in each condition. We report proportion of 'yes' responses rather than the actual number of responses to facilitate comparison with Leslie et al.'s (2011) results.

Table 1. Mean Proportion (SE) of 'yes' responses as a function of context and determiner type in Exp. 1. The subtractions between the relevant conditions are also given.

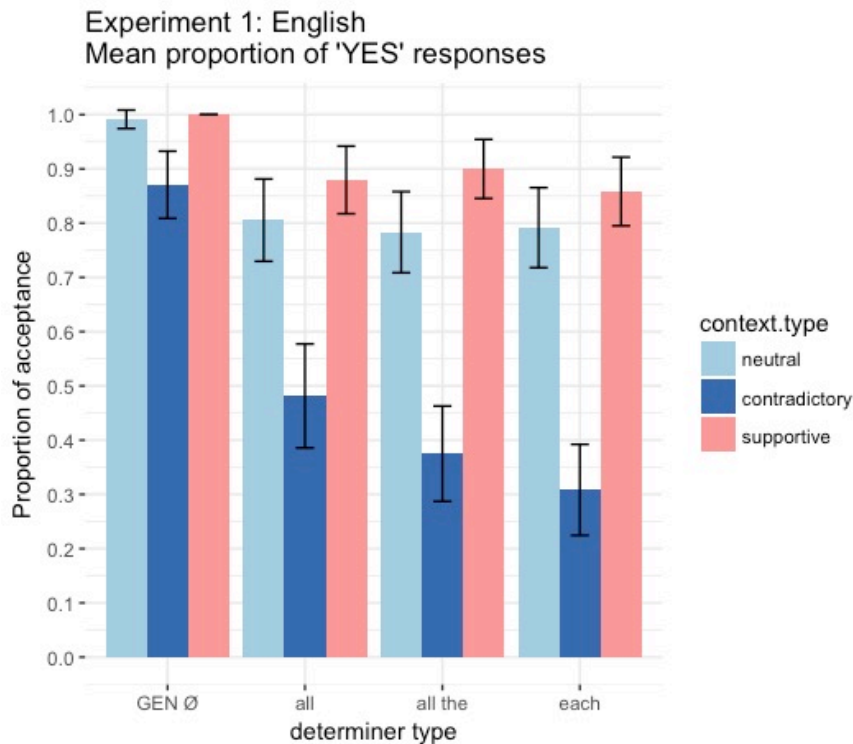| Condition	Neutral	Contradictory	Supportive	Contradictory-Neutral	Supportive-Neutral
GEN (ø)	99.14 (3.12)	87.07 (0.86)	100 (0)	-12	1
all	80.56 (3.82)	48.15 (4.83)	87.96 (3.14)	-33	7
all the	78.33 (3.78)	37.50 (4.43)	90 (2.76)	-40	12
each	79.17 (3.72)	30.83 (4.23)	85.83 (3.2)	-48	7

Figure 1. Mean proportion of 'yes' responses as a function of context and determiner type. Error bars represent 95% confidence intervals.

As we see above, overall, generics were accepted at higher rates than universals, as expected, given that we had chosen items that were true in generic form. Both in the neutral and the supportive condition acceptance rates for generics were at ceiling (99% and 100% respectively) and were only lower in the contradictory condition (87%). With universals, the picture is more complicated. All three universals ('all', 'all the', 'each') were accepted at similar rates in both the neutral and the supportive condition, showing only a small increase in the supportive condition. In the neutral condition, 'all'-statements were accepted 81% of the time, 'all the'-statements 78% of the time and 'each'-statements 79% of the time. In the supportive condition, 'all'-statements were accepted 88% of the time, 'all the'-statements 90% of the time and 'each'-statements 86% of the time. Universals after a contradictory context yielded fewer acceptances overall, as expected: 'all'-statements were accepted at 48%, 'all the'-statements at 38% and 'each'-statements at 31%.

We used R (R Core Team, 2016) and the lme4 package (Bates, Maechler, Bolker & Walker, 2015) to perform a generalised mixed-effects linear analysis of the effects of between determiner and context on the yes/no response, specifying a binomial family. Responses were treated as a dummy coded categorical variable and were modelled with

glmer. First, we fitted a full model with det.type and context.type as fixed effects (with an interaction term) and with random intercepts for subjects and items. We performed a likelihood ratio test of the full model with an interaction term against a model without the interaction term and the comparison proved non-significant ($\chi^2(6) = 8.3455$, $p = .214$). Including an interaction term did not significantly improve model fit, so we used the model without the interaction term for all subsequent analyses/comparisons.

We then fitted versions of the full model, in which a single effect was removed and we compared the reduced model to the model without interaction. To test the main effect of context, we removed context. A likelihood ratio test of the model without interaction against the model without context proved significant ($\chi^2(2) = 311.81$, $p < .001$). Thus, we concluded that there was a main effect of context. To test the main effect of determiner, we removed determiner. A likelihood ratio test of the model without interaction against the model without determiner proved significant ($\chi^2(3) = 58.183$, $p < .001$). Thus, we concluded that there was a main effect of determiner.

As we see in table 2, in the neutral condition, the difference between generics and all universal quantifiers was statistically significant (all ps $< .001$). Generics in the neutral condition are significantly different from generics in the contradictory condition ($p = .004$), but not significantly different from the supportive condition ($p = .76$). Thus, manipulating context seems to influence how people interpret generics too, to a certain extent.

Table 2. Estimates, standard errors, z values and p values of the full glmm in Exp. 1.

	Estimate	Std. Error	z value	Pr(>|z|)
Intercept (generic, neutral)	5.8432	1.1150	5.241	1.6e-07 ***
det (all)	-3.9450	1.1708	-3.369	0.000753 ***
det (all the)	-4.0789	1.1615	-3.512	0.000445 ***
det (each)	-3.9228	1.1636	-3.371	0.000748 ***
context (contradictory)	-3.1623	1.0847	-2.915	0.003553 **
context (supportive)	13.7194	44.9064	0.306	0.759978
det x context (all, contradictory)	1.1622	1.1434	1.016	0.309411
det x context (all the, contradictory)	0.6892	1.1392	0.605	0.545210
det x context (each, contradictory)	0.0946	1.1473	0.082	0.934285
det x context (all, supportive)	-13.0339	44.9071	-0.290	0.771631
det x context (all the, supportive)	-12.4670	44.9084	-0.278	0.781312
det x context (each, supportive)	-13.1128	44.9067	-0.292	0.770285

Moving now to the universal quantifiers. On the surface, we do obtain many 'yes' responses to universal quantifiers, as in the generics-as-default literature. We predicted that, across the board, contradictory context should decrease acceptances, while supportive context should increase them. More importantly, we had specific predictions about the relative rates between the universal quantifiers depending on their sensitivity to QDR, which the default generic bias cannot predict. In order to appreciate the relative effect of context on acceptance rates, we subtracted the average means of the neutral condition from the average means of the contradictory and supporting conditions. We interpreted the rates obtained as the relative effect of context on acceptance rates plotted in Fig. 2 below.
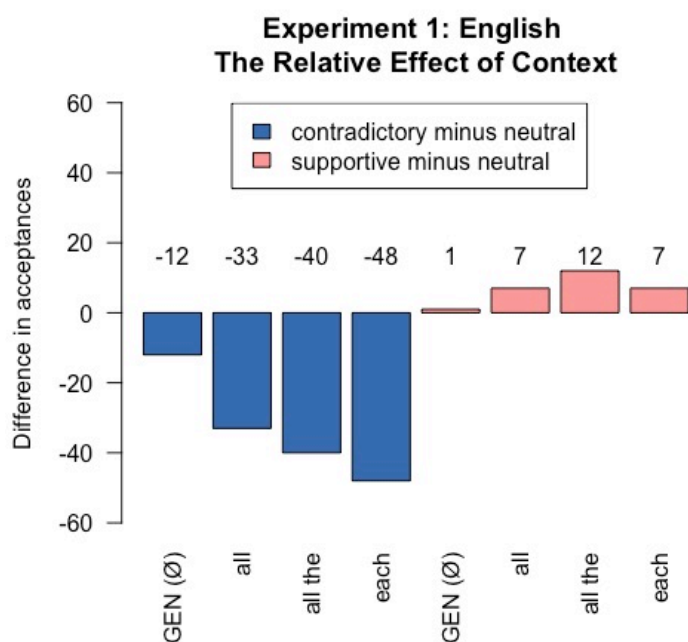

Figure 2. The relative effect of context in Exp. 1. Acceptance rates in the neutral condition are plotted with the baseline condition as 0. Negative values show the subtracted average means of the neutral condition from the average means of the contradictory condition and positive values show the subtracted average means of the neutral condition from the average means of the supportive condition.

In a second planned analysis, seen in table 3, we were interested in planned comparisons between the different universal quantifiers. We built a model only with the universal quantifiers, 'all', 'all the' and 'each'. We first looked at effects related to 'all' and the three levels of context. 'All' in the neutral context differed significantly from 'all' in the contradictory context ($p < .001$), but not from 'all' in the supportive context ($p = .112$). We then looked at effects related to 'all the' and the three levels of context. 'All the' in the

neutral context differed significantly both from 'all the' in the contradictory context ($p < .001$), and 'all the' in the supportive context ($p = .004$). 'Each' in the neutral condition differed significantly from 'each' in the contradictory context ($p < .001$), but not from 'each' in the supportive context ($p = .145$). In the neutral condition, there were no significant differences between 'all' and 'all the' ($p = .809$) or between 'all' and 'each' ($p = .961$). In the contradictory condition, 'all' differed significantly from 'each' ($p = .038$), but not from 'all the' ($p = .221$). In the supportive condition there were no significant differences between 'all' and 'each' ($p = .918$) nor between 'all' and 'all the' ($p = .488$).

Table 3. Estimates, standard errors, z values and p values of the glmm without generics in Exp. 1.

	Estimate	Std. Error	z value	Pr(>|z|)
Intercept (all, neutral)	1.90580	0.42752	4.458	8.28e-06 ***
det (all the)	-0.13329	0.55003	-0.242	0.8085
det (each)	0.02695	0.55556	-0.049	0.9613
context (contradictory)	-2.00459	0.37130	-5.399	6.70e-08 ***
context (supportive)	0.68454	0.43037	1.591	0.1117
det x context (all the, contradictory)	-0.47904	0.51241	-0.935	0.3499
det x context (each, contradictory)	-1.07973	0.52983	-2.038	0.0416 *
det x context (all the, supportive)	0.57050	0.61347	0.930	0.3524
det x context (each, supportive)	-0.08964	0.59212	-0.151	0.8797

Thus, we obtain the effect for the contradictory context exactly as predicted for the universal quantifiers. The relative effect is bigger for those quantifiers that require QDR because of their semantics ('all the', 'each') than for the one that allows but does not require it/may resist it ('all'). The prediction about the supportive context was not borne out except for 'all the', where we do see a significant increase compared to the neutral condition. Adding explicit information supporting the statement hardly mattered, as acceptances did not rise significantly (increase of 1% for generics, 7% for 'all', 7% for 'each'), except for 'all the' at 12%. The high acceptance rates in the neutral condition are possibly due to participants being charitable and/or exceptions not being immediately salient. The atypical nature of 'all' does not seem to be a likely explanation of the GOG effect given that the differences in acceptances across quantifiers in the neutral condition were minimal, that is, the difference between 'all' and 'all the' was 3% and the difference between 'all' and 'each' was 2%.

Regarding our control items, across participants, acceptance rates for definitional statements were almost at ceiling, at 97%, while false generalisations were rejected most of

the time with acceptance rates at 31%. These results were as expected and are in line with previously discussed rates in the literature (e.g. Leslie et al., 2011, report 90% and 38% acceptance respectively in their experiment 1).

An outstanding issue is why there were any acceptances in the contradictory condition for 'all the' and 'each' given the prediction that participants have to necessarily restrict their interpretation to the counterexamples given in the discourse. These acceptances (at 38% for 'all the' and at 31% for 'each') could still be interpreted by proponents of the generics-as-default view as instances of a GOG effect. We believe though, that in order to conclude that there is such an effect, one must not only point out that there are still acceptances that are not accounted for by domain restriction, but also to rule out other already existing alternative explanations too. Invoking the other factors that might influence 'all' is not a viable option for 'all the' or 'each', because the subkind interpretation is illicit and their meaning is not prone to hyperbolic/loose interpretation. Some of these acceptances might be potentially attributed to the strong association between the concept TIGER and the concept STRIPE (see footnote 3). This though would also mean that participants were not actually parsing the statements as a whole within a context, but they were only focusing on the lexical meaning of the words involved, which is highly unlikely given the differences between conditions.

Other factors that contribute to the non-zero acceptances are outlined here: (a) participants' acceptances of false fillers such as 'London is in Egypt', which were non-quantified statements concerning factual knowledge, were at 7% suggesting that a new baseline should be drawn treating 7% of acceptances across the board as pure errors (e.g. due to participants being inattentive, pressing the incorrect key, etc.), that is, if 7% of these acceptances were errors, we are left with 31% for 'all the' and 24% for 'each' to be accounted for, (b) participants' acceptances of false, non-quantified generalisations such as 'animals are reptiles' or 'books are paperbacks' at 31% further indicates that people generally find judging these kind of statements difficult. This suggests again that the actual baseline error rate is higher.

Two independent phenomena, quite broadly attested in the literature, might have also influenced these acceptances: (a) the *acquiescence bias*, i.e. the tendency to endorse an assertion made in a question regardless of the assertion's content (Krosnick & Presser, 2010), and (b) participants are known to be pragmatically charitable (Grice, 1975), thus they would try to find an interpretation that would render the statement 'true' even despite the salient context (see for instance Chemla and Bott, 2013, for the relevance of charitable interpretation in pragmatic reasoning).

25

Discussion of Experiment 1

We set out to explore one of the alternative explanations for the judgement data that concern universally quantified statements, which have been used as evidence of a GOG effect. The present study provides experimental evidence for the relevance of a QDR-based explanation for much of the attested behaviour. In our study, context did not only affect acceptance rates for 'all', 'all the', 'each', but it further predicted the levels of QDR depending on the level of context. The effect of context was greater for 'all the' and 'each', two quantifiers that require QDR, while it was smaller for 'all', whose domain is only optionally restricted and may resist it. This leads us to conclude that Leslie et al. (2011) prematurely understated the relevance of context as a principled way of explaining the attested behaviour. Furthermore, even though they only discuss 'all' they make general claims about (universal) quantification being prone to the GOG effect. We argue that drawing conclusions about universal quantification (and by extension about genericity) requires more subtle manipulations. The differences we found between the different universal quantifiers are predicted according to the QDR view advanced here. We also find that manipulating context seems to affect generics to some extent too. This seems contrary to the received view that takes generics to strictly resist contextual narrowing (Krifka, 1987) but is in line with recent work arguing that generics display some context sensitivity (see Sterken, 2015, and references therein). An explanation for this result is offered in the general discussion (section 7).

6.   Experiment 2: Generalisations in Greek

With respect to generics in Greek, as discussed in section 4, a statement like *i tighris ehun righes* 'DET tigers have stripes' is potentially ambiguous between a generic and a definite (anaphoric) interpretation. Given this ambiguity, one might expect fewer acceptances of 'generics' in Greek across the board. In the neutral and supportive contexts both interpretations would give rise to the same target response. Judging *i tighris ehun righes* after a neutral context, one should accept it in both cases: (a) under a generic interpretation, they should judge it as 'true' given that they know that this is a characteristic of tigers that holds in general, and (b) under an anaphoric interpretation, they should judge it as 'true' given that they have no reason to assume that the specific tigers mentioned in the discourse are not normal instances of tigers. Judging *i tighris ehun righes* after a supportive context, one is

expected to accept it in both cases too: (a) under a generic interpretation, for the same reason as above, and (b) under an anaphoric interpretation, they should judge it as 'true' given that the context provides them with positive information about the property in question. Because the supportive context explicitly licences the anaphoric interpretation, we expect higher acceptance rates in this context compared to the neutral. The expected response only differs in the case of the contradictory context depending on what interpretation one assigns to the statement: (a) under a generic interpretation, one is expected to accept it given that generics tolerate exceptions, but (b) under an anaphoric interpretation, one is expected to reject it given that the individuals introduced in the discourse lack the relevant property. In the contradictory context, we expect lower acceptances than both neutral and supportive contexts.

A related prediction concerns the issue of whether there is a preference for the generic over the definite (anaphoric) interpretation in Greek. Previous research on Spanish definite plural generics (Ionin and Montrul, 2010), which, like Greek, are ambiguous between a generic and an anaphoric interpretation, finds that Spanish native speakers were far more likely to interpret such statements as generic (81%) than anaphoric after a context that was similar to our contradictory context, irrespective of whether it made the statement 'true' or 'false'. Regardless of whether that preference was a task-related effect, a possibility the authors acknowledge, it is worth-while investigating whether such a bias exists in Greek too.

With respect to universals, according to the generics-as-default view, given that generics are the default way children and adults generalise, the tendency to overgeneralise when it comes to universals ought to be universal. Thus, speakers of languages other than English might also be prone to make erroneous judgments concerning universal generalisations under some circumstances.

Cross-linguistic evidence can additionally tackle the issue of whether the effect is due to the atypical behaviour of the English quantifier 'all'. Although *oli i* contains the definite article, this quantifier behaves similarly to the English quantifier 'all' and not to 'all the' (see Tsili, 2001; Tsamadou-Jacoberger, 2006; Giannakidou, 2012). For instance, *oli i* can be used to refer to either all students in general (*oli i mathites latrevun tis diakopes* 'all students love holidays') or it can be contextually restricted to a specific set of students (*oli i mathites hamoghelusan* 'all the students were smiling'). Thus, Greek differs from English in that it does not have two versions of the quantifier 'all'. The Greek quantifier *oli i* optionally allows restriction. Participants who accept a statement like 'all tigers have stripes' in Greek might be applying QDR to the quantifier. However, as is the case with 'all', there are other alternative

explanations which might lead someone to accept a statement with *oli i*: (a) ignorance and/or (b) the subkind interpretation. The generics-as-default view, in contrast, has no principled way of explaining any potential differences in acceptance rates depending on the level of context (neutral, contradictory, supportive) for *oli i* 'lit. all the' any more than it does for 'all'. We will be glossing *oli i* as 'all' for the remainder of the paper.

6.1 Method

6.1.1   Participants and procedure

60 volunteers (47 female; aged 19-71; mean age 36.28; SD 16.08) participated in the experiment over the Internet. Participants were recruited through mailing lists and word of mouth. All spoke Greek as their first language and provided their informed consent to enter a prize draw. The study followed the same guidelines and used the same procedure as Experiment 1, except that it was in Greek and only compared generics (definite plurals) to one type of universal quantifier (*oli i* 'all').

6.2 Materials and design

Participants judged the same set of 84 statements as in Experiment 1, including 48 fillers presented in a randomized order. The 12 test items of majority characteristic statements were exactly the same. The 24 control items (12 definitional statements and 12 false generalisations) were the same except for four items, which were replaced in order to sound more natural in Greek. Given the potential ambiguity between an anaphoric and a generic interpretation of a definite plural, the contexts for the false generalisations were modified in order to bias participants towards a generic interpretation. As in the English experiment, all the contexts and items were normed beforehand by Greek native speakers, who did not take part in the experiment. The two conditions we manipulated for the majority characteristic items were:

a.   determiner type: definite plural generic/*oli i* 'all'
b.   context type: neutral/contradictory/supportive

Given the 4 determiners (generic/*all/all the/each*) we created 2 lists with 3 sublists each that varied with respect to the pairing of the items with context type, which gave us 6 lists in total. 10 participants were assigned randomly to each of the 6 lists.

6.3 Results and discussion

The final analysis included 57 participants. Three participants were excluded as they responded correctly to fewer than 10 out of the 12 definitional statements.

Table 4 summarises the proportion of 'yes' responses to the TVJ question for the test items (majority characteristic statements) in each condition and Fig. 3 plots these results graphically.

Table 4. Mean Proportion (SE) of 'yes' responses as a function of context and determiner type in Exp. 2. The subtractions between the relevant conditions are also given.

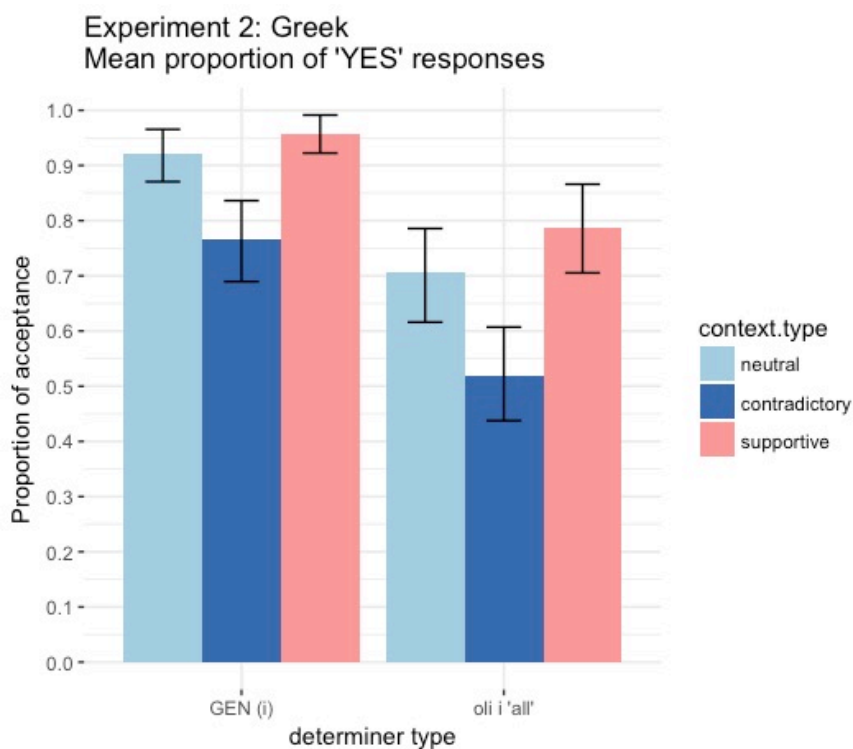| Condition	Neutral	Contradictory	Supportive	Contradictory-Neutral	Supportive-Neutral
GEN (i)	92.24 (2.49)	76.72 (3.94)	95.69 (1.89)	-15	4
oli i	70.54 (4.33)	51.79 (4.74)	78.57 (3.89)	-19	8

Figure 3. Mean Proportion of 'yes' responses as a function of context and determiner type. Error bars represent 95% confidence intervals.

29

As we see above, generics were accepted at higher rates than universally quantified statements with *oli i* 'all' across the board. This was as expected both for the neutral and the supportive conditions given that both possible interpretations of the nominal (generic/anaphoric) render the statements 'true'. Thus, generics were judged as true at very high rates (92% in the neutral and 96% in the supportive condition). Acceptance was lower in the contradictory condition, at 77%. This was expected, given that one of the interpretations (the anaphoric one) made the statement false. Even though the anaphoric interpretation was the most salient one for this task, our findings seem to support a strong preference for the generic reading at rates similar to the ones observed for Spanish (Ionin and Montrul, 2010). Universals with *oli i* 'all' were accepted at 71% in the neutral condition and at 79% in the supportive condition. Acceptances in the contradictory condition were at 52%. These differences in acceptances of the universally quantified statements show that context is a major licensor of the attested behaviour.

We followed the same procedure we used for English to analyse the Greek data. First, we fitted the full model with det.type and context.type as fixed effects (with an interaction term) and subject and item.no as random effects – see table 6 below. After building the full model, we built a model with the same fixed and random effects, but without an interaction term. We performed a likelihood ratio test of the full model with an interaction term against the model without the interaction term and the comparison proved non-significant ($\chi^2(2) =$ 1.343, $p = .51$). Thus, adding an interaction term did not significantly improve model fit, so we used the model without the interaction term for all subsequent analyses/comparisons.

We then fitted versions of the full model, in which a single effect was removed and we then compared the reduced model to the model without interaction. In order to test the main effect of context, we removed context.type. We performed a likelihood ratio test of the model without interaction against the model without context.type and the comparison proved significant ($\chi^2(2) = 58.824$, $p < .001$). Thus, we concluded that there was a main effect of context. In order to test the main effect of determiner, we removed det.type. We performed a likelihood ratio test of the model without interaction against the model without det.type and the comparison proved significant ($\chi^2(1) = 17.544$, $p < .001$). Thus, we concluded that there was a main effect of determiner.

As we see in table 5, focusing on the neutral condition, generics differ significantly from *oli i* ($p < .001$). Focusing on generics in different levels of context, the difference between the generic neutral is significantly different from the generic contradictory ($p <$ .001), but not between the generic neutral and the generic supportive ($p = .11$). Focusing on

the universal *oli i* and the different levels of context, the difference between *oli i* neutral and *oli i* contradictory is significant ($p < .001$), while the difference between *oli i* neutral and *oli i* supportive is not significant ($p = .055$).

Table 5. Estimates, standard errors, *z* values and *p* values of the full glmm in Exp. 2.

	Estimate	Std. Error	z value	Pr(>|z|)
Intercept (generic, neutral)	3.7630	0.6698	5.618	1.93e-08 ***
det (oli i)	-2.3895	0.6752	-3.539	0.000401 ***
context (contradictory)	-1.7637	0.5090	-3.465	0.000530 ***
context (supportive)	1.1032	0.6897	1.599	0.109721
det x context (oli i, contradictory)	0.4620	0.6226	0.742	0.458100
det x context (oli i, supportive)	-0.3217	0.7943	-0.405	0.685498

For universals, we obtained many 'yes' responses to *oli i* 'all' statements ranging from 79% (supportive) to 71% (neutral) and 52% (contradictory). These data would most probably be attributed to a GOG effect by proponents of the generics-as-default view. Nevertheless, we believe that it is important to focus first on the quantifier's sensitivity to QDR. In the Greek design we did not have any quantifier that required QDR. *Oli i* 'all' does not require QDR, but optionally allows it. Plotting in Fig. 4 below the relative effect of context we note that the differences between the different levels of context for *oli i* 'all' are in the direction expected given the results of Experiment 1, with fewer acceptances in the contradictory condition and more acceptances in the supportive condition. The participants who accept a statement like 'all tigers have stripes' in Greek might have done so because of the following reasons: (a) the available option for QDR for *oli i* 'all', (b) because they used a subkind interpretation or (c) because they were ignorant of the facts.

Experiment 2: Greek
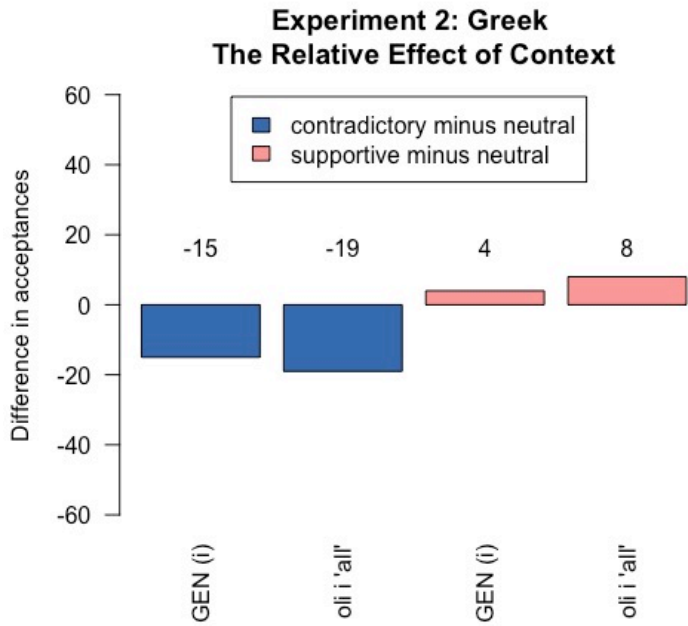**The Relative Effect of Context**

Figure 4. The relative effect of context in Exp. 2. Acceptance rates in the neutral condition are plotted with the baseline condition as 0. Negative values show the subtracted average means of the neutral condition from the average means of the contradictory condition and positive values the subtracted average means of the neutral condition from the average means of the supportive condition.

We observe an interesting difference if we compare the English 'all' results to the Greek *oli i* 'all' results. The relative effect of context in the contradictory condition seems bigger in English, with a decrease of 33% compared to the neutral condition, than in Greek, where the decrease is 19%. Focusing on the raw scores though reveals that acceptance rates are similar in both languages in the contradictory condition (at 48% for English and 52% for Greek), but there are far fewer acceptances for *oli i* in the neutral condition (71%) than for *all* (81%), leading to a smaller relative effect for the context manipulation. We could attribute this to the presence of the control items, which used definite plurals in Greek, which are potentially ambiguous in Greek but not in English, as discussed above. Greek participants might have been more aware of multiple interpretations of a given statement than their English counterparts, even in the neutral condition, and thus less prone to accept the statements.

Moreover, Greek generics are slightly more affected by the context manipulation than English generics, as expected given the potential ambiguity in Greek. Acceptances in the neutral condition are less common in Greek (92%) compared to English (99%) and the

relative effect of context in the contradictory condition is bigger in Greek (a decrease of 15%) than in English (a decrease of 12%).

Regarding our control items, across participants definitional statements were accepted 96% of the time and false generalisations were accepted 7% of the time. The contexts for false generalisations were different from the English ones (where acceptances were at 31%), as the context in Greek was biased towards a generic interpretation (in order to rule out the anaphoric interpretation of the statement). This might have made other subkinds of the relevant kind more salient and so rejecting the statement was facilitated for Greek participants.

To sum up, experiment 2 addressed one interpretation of the unmarkedness claim with respect to generics and showed that the picture is more complicated as soon as one moves from English to a language like Greek with a different configuration of nominals involved in genericity, quantification and definiteness. The results confirmed that our context manipulation works in two different languages and that QDR is a viable explanation for much of the purported GOG effect cross-linguistically.

7. General Discussion

We set out to explore the scope of the purported generic overgeneralisation (GOG) effect, as it has been proposed as one of the main pieces of evidence for the generics-as-default view (Leslie 2007, 2008; Gelman, 2010; Leslie et al. 2011). According to its proponents, this effect concerns participants accepting universally quantified statements such as 'all ducks lay eggs' and 'all tigers have stripes' as true, even though they should reject them as false. The explanation given for the pattern of judgment data observed is that participants fail to properly evaluate the universal statements in question and, erroneously, rely on their judgments of the corresponding generics, in this case 'ducks lay eggs' and 'tigers have stripes', which are true and tolerant of exceptions.

The studies presented here show that attributing the attested behaviour to a default overgeneralisation bias such as the GOG effect is not yet fully warranted. We argue that Leslie et al. (2011) ruled out alternative explanations for these data prematurely and we furthermore addressed the size of one of the alternative explanations for the attested behaviour. Our studies provide experimental evidence for the relevance of a QDR-based explanation of much of the behaviour attributed to the GOG effect. In our studies, we focused on majority characteristic statements like 'all tigers have stripes' and manipulated context in

order to show that what is at play is not a generic bias, but people judging the generalisations largely based on what they know about the semantics and pragmatics of genericity and quantification and on independently motivated mechanisms related to these interpretations such as QDR. We used three levels of context: (a) neutral, where the information in the context does not interact with the truth value of the critical statement; (b) contradictory, where an exception which should rule out a universally quantified statement is made salient, and (c) supportive, where a paraphrase of the critical property is given, which makes its generality salient. Based on these three levels of context and the sensitivity to QDR that the universal quantifier used in the critical statement shows, we were able to predict much of the variation in acceptance rates. The context affected acceptance rates for 'all' and *oli i* 'all' and other universal quantifiers ('all the', 'each') and thus it predicted the levels of QDR. Our data show that drawing conclusions about universal quantification (and by extension about genericity) requires more subtle manipulations. The differences we found between the different universal quantifiers are predicted according to the QDR view advanced here but are inconsistent with the generics-as-default view, as we understand it. According to the generics-as-default view, the 'more marked' System 2 universal statements should be 'erroneously' interpreted as generic statements at similar rates, as the attested behaviour is due to an underlying generic bias that involves incorrect processing of the quantifiers and does not depend on the properties of the quantifiers themselves.

The claim that generics are unmarked compared to quantifiers has also played an important role in the generics-as-default proposal. We argued that the claim is not well-substantiated for any language and that without a proper definition of markedness, this claim is impossible to evaluate. We made a first step towards that direction by looking at a language where generics are expressed differently than in English. Crucially, for the issue at hand, if we assume a definition of markedness as formal markedness, i.e. surface level overt realisation, following Haspelmath (2006), Greek generics are not less marked that quantifiers. Our results show that the QDR view advanced here correctly predicts the pattern of responses depending on the type of context and the sensitivity of *oli i* 'all' to QDR, whereas the generics-as-default view has no principled way of explaining the variable acceptance rates between contexts.

A fact that bears further discussion is what looks like context effects for generics in English. This seems contrary to the received view that takes generics to strictly resist contextual narrowing (Krifka, 1987) but is in line with recent work (Sterken, 2015) that argues that generics may be subject to contextual restriction after all. Looking more closely at

the examples Sterken (2015) discusses reveals that generics' context sensitivity might be tied only to some specific cases and not to all types of generic statements. Sterken's examples highly depend on the context of utterance and the question under discussion e.g. the statement 'dobermans have floppy ears' is true in the context of evolutionary biology, but false in the context of dog breeding. Our materials manipulated context in the sense of making some exceptions salient or not, but we did not introduce different contexts of utterance in the sense described above. Thus, it is not clear that our materials constitute cases of context sensitivity for generics in the sense of Sterken. There are, however, other alternative explanations for the observed behaviour. It might be that the contradictory contexts with the explicit mention of counterexamples made the participants doubt their beliefs with respect to the prevalence of the property in question. Thus, some participants might have (wrongly) updated their belief about the prevalence of the property in order to align with what was given to them in the context of the experiment leading them to reject (true) generic statements in English.

A further issue to be addressed in future studies is the relevance of the other alternative explanations for the attested behaviour. Even though we have shown that QDR is a major contributing factor in the attested behaviour, the other ways via which people might accept the statements in question still require investigation. In this respect, the subkind interpretation seems particularly relevant and we have argued above (see section 1.2b above) that the way this has been addressed in previous literature is not satisfactory. It is important to note that all the statements used here and in related studies are about animals, which belong to natural kinds. Natural kinds lend themselves easily to a subkind interpretation, as people know that animals have different subspecies and belong to a well-described biological classification (see e.g. the Linnaean taxonomy). It is easy to imagine different subkinds of tigers that all belong to the kind tiger, which is characterised by the property of having stripes. Or that different subkinds of ducks all have in common the fact that they lay eggs. A future study could address the relevance of the subkind interpretation, as it is interesting to understand the degree to which these phenomena are tied to animal kinds, for which a characterisation at species level is readily available alongside different subkinds. In table 6 we summarise the alternative explanations to the GOG effect ordered by effect size.

Table 6. Alternative explanations to the GOG effect ordered by effect size.

Explanation	Study	Quantifier	Effect Size		Comment
1. **Quantifier Domain Restriction**	Our experiments		majority		
	English	'all'	**-33%**		Difference in acceptances
	English	'all the'	**-40%**		between the neutral and the
	English	'each'	**-48%**		contradictory condition
	Greek	'oli i'	**-19%**		
	Leslie et al. (2011)	'all'	majority	minority	
			-18%	**-21%**	Difference in acceptances between exp. 1 and 2a
	Leslie et al. (2011)	'all'	**-1.6%**		Paraphrases including subset restriction in exp. 2b
2. **Ignorance**	Leslie et al. (2011)		majority	minority	
		'all'	**+20%**[7]	**-19%**	Difference in acceptances between exp. 3a and 3b
3. **Error**	Our experiments				
	English	no Q	**6.67%**		Acceptances of false fillers
	Greek	no Q	**5.45%**		
4. **Atypical 'all'**	Our experiments	'all the' vs. 'all'	**-3%**		Acceptances in the neutral condition
		'each' vs. 'all'	**-2%**		
5. **Subkind interpretation**	Leslie et al. (2011)	'all'	**-1%**		Paraphrases including subtypes in exp. 2b

If each of these factors contributes independently and cumulatively to the 'erroneous' acceptance rates for universally quantified generalisations characterised as GOG errors in the literature, the entirety of the 78% acceptance rate for universally quantified majority

---

[7] See Leslie et al. (2011) and discussion in Lazaridou-Chatzigoga, Stockall, & Katsos (2017) for consideration of this surprising *increase* in the rates at which statements such as 'all tigers have stripes' were accepted as true when the participants performed the knowledge task before evaluating the statements. It is important to note that the knowledge task failed to test whether participants are aware that majority characteristic statements also have exceptions, that is, participants in the knowledge task were not asked to judge whether 'albino tigers have stripes' or 'amputated horses have four legs', but were only asked to judge false minority characteristic statements such as 'male ducks lay eggs'.

characteristic statements in Leslie et al.'s (2011) experiment 1 could be explained by these factors, with hardly any residue requiring a GOG explanation. Our key claim is that such alternative explanations need to be seriously entertained, and systematically evaluated, before invoking a GOG effect. In the future, a careful experimental examination of all the other known semantic and pragmatic factors that independently lead to behaviour similar to the GOG effect will help to properly evaluate the contribution of the GOG effect on the interpretation of universal statements.

Another crucial aspect that deserves closer attention in future studies is acceptance rates per item. The previous literature does not report these, but it is important to note that in our studies we do find that some items are accepted/rejected more easily than others. For example, 'all butterflies have wings' is accepted 75% of the time in the contradictory context, despite the context specifically discussing injured, wingless butterflies, while 'all leopards have spots' is only accepted 11% of the time in this condition. Similar item specific variation is also found for the other quantifiers, and in the Greek data set. It is not obvious what could be driving such differences.

In sum, we argue that two central assumptions/ingredients of the generics-as-default view, the generic overgeneralisation effect and the unmarkedness claim, have to be scrutinised in studies that manipulate additional factors and to be tested cross-linguistically. The general thrust of this work is that, rather than being under the influence of a blanket default bias that covers all cases across the board, adult participants who make these judgements are largely sensitive to the subtle interplay of quantifier semantics and pragmatics on the one hand, and context on the other. This approach has the advantage of accounting for big portions of the attested data by means of an independently motivated factor, QDR, without the need to postulate a new mechanism such as the GOG effect which is specifically at play exclusively for universal generalisations.

The present data serve as the first demonstration of how an interdisciplinary programme that brings systematicity and consistency to the study of generics and universals across a range of diverse domains of inquiry can significantly advance our understanding of how we make generalisations and how we express them in language. They also suggest more directions for future investigation. By broadening the investigation of generic and universal generalisations to languages other than English, we can clarify the language-specific and language-general features of generics. It will also be important to address the nominal system in the language under investigation more broadly, as the interplay of genericity, quantification and definiteness seems to have repercussions for the availability of certain

interpretations. Finally, the present research raises the question of what factors more generally influence the way people express and understand generalisations by drawing special attention to pragmatic reasoning and linguistic knowledge of the system of the language in question.

Acknowledgments

References

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67(1)*, 1-48.

Beghelli, F., & Stowell, T. (1997). Distributivity and negation. In A. Szabolcsi (Ed.), *Ways of Scope Taking* (pp. 71-107). Dordrecht: Kluwer.

Behrens, L. (2000). *Typological Parameters of Genericity*. Arbeitspapier 37 (Neue Folge). Institut für Sprachwissenschaft, Köln: Universität zu Köln.

Behrens, L. (2005). Genericity from a Cross-linguistic Perspective. *Linguistics, 43(2)*, 275-344.

Carlson, G. (1977). Reference to kinds in English. PhD dissertation, University of Massachusetts, Amherst.

Chemla, E. & Bott, L. (2013) Processing presuppositions: Dynamic semantics vs pragmatic enrichment, *Language and Cognitive Processes*, *28:3*, 241-260.

Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics, 6*, 339-405.

Chomsky, N. (2000). New Horizons in the Study of Language and Mind. Cambridge: Cambridge University Press.

Claridge, C. (2011). Hyperbole in English: A corpus-based study of exaggeration. Cambridge: Cambridge University Press.

Cohen, A. (2001). Relative Readings of 'many', 'often' and Generics. *Natural Language Semantics*, *9*, 41-67.

Collins, A. M. & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behaviour*, 8, 240-248.

Dahl, Ö. (1975). On generics. In E. Keenan (Ed.) *Formal semantics of natural language*, (pp. 99-111). Cambridge: Cambridge University Press.

Dahl, Ö. (1985). *Tense and Aspect Systems*. Oxford: Blackwell.

Dahl, Ö. (1995). The marking of the episodic/generic distinction in tense-aspect systems. In Krifka et al. (1995), (pp. 412-425). Chicago: Chicago University Press.

Dayal, V. (2004). Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy, 27*, 393-450.

Dienes, Z, & Perner, J. (2002). What sort of representation is conscious? *Behavioral and Brain Sciences, 25*, 336-337.

von Fintel, K. (1994). Restrictions of quantifier domains. PhD dissertation: University of Massachussets, Amherst.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19(4)*, 25-42.

Gelman, S. A. (2010). Generics as a window onto young children's concepts. In F.J. Pelletier (Ed.), *Kinds, things, and stuff: The cognitive side of generics and mass terms*. New directions in cognitive science (pp. 100-123). New York: Oxford University Press.

Giannakidou, A. (2012). The landscape of Greek quantifiers. In E. L. Keenan and D. Paperno (Eds.), *Handbook of Quantifiers in Natural Language*, Studies in Linguistics and

Philosophy. Volume 90 (pp. 285-346). Dordrecht: Springer.

Giannakidou, A., & Stavrou, M. (1999). Nominalization and ellipsis in the Greek DP. *The Linguistic Review, 16*, 295-332.

Gil, D. (1995). Universal Quantifiers and Distributivity. In E. Bach, E. Jelinek, A. Kratzer and B. Partee (Eds.), *Quantification in Natural Languages* (pp. 321-362). Dordrecht: Kluwer.

Grice, H. P. (1975). Method in Philosophical Psychology: From the Banal to the Bizarre. In *Proceedings and Addresses of the American Philosophical Association* (pp. 23-53). Oxford: Clarendon Press.

Hampton, J. A. (2015). Categories, prototypes and exemplars. In N. Reimer (Ed.), *Routledge Handbook of Semantics.* Abingdon: Routledge.

Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics, 42*, 25-70.

Heim, I. (1991). Artikel und Definitheit [Articles and Definiteness]. In A. von Stechow & D. Wunderlich (Eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung [An International Handbook of Contemporary Research]* (pp. 487-535). Berlin: de Gruyter.

Hollander, M. A., Gelman, S. A., & Star, J. (2002). Children's Interpretation of Generic Noun Phrases. *Developmental Psychology, 36(6)*, 883-894.

Ionin, T. and Montrul, S. 2010. The role of L1-transfer in the interpretation of articles with definite plurals in L2-English. *Language Learning*, 60, 4, 877-925.

Jönsson, M. L., & Hampton, J. A. (2006). The inverse conjunction fallacy. *Journal of Memory and Language, 55*, 317-334.

Kahneman, D. & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (pp. 49-81). New York: Cambridge University Press.

Khemlani, S., Leslie, S.J., Glucksberg, S., & Rubio-Fernandez, P. (2007). Do ducks lay eggs? How humans interpret generic assertions. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th annual conference of the Cognitive Science Society* (pp. 395-400). Nashville, TN: Cognitive Science Society.

Kratzer, A. (2004). Covert quantifier restrictions in natural languages. Talk given at Palazzo Feltrinelli in Gargnano June 11, 2004.

Krifka, M. (1987). An Outline of Genericity, partly in collaboration with C. Gerstner.

Technical Report SNS-Bericht 87-25 Seminar für natürliche Systeme, Tübingen UniversityGermany.

Krifka, M., Pelletier, F., Carlson, G., ter Meulen, A., Chierchia, G. & Link, G. (1995). Genericity: An Introduction. In G. Carlson and F. J. Pelletier (Eds.) *The Generic Book* (pp. 1-125). Chicago. Chicago University Press.

Krosnick, J. A., & Presser, S. (2010). Question and Questionnaire Design. In J. D. Wright & P. V. Marsden (Eds.), *Handbook of Survey Research* (Second Edition) (pp. 263-313). West Yorkshire, England: Emerald Group.

Lazaridou-Chatzigoga, D. (2009). Greek Generic Noun Phrases Involving the Free Choice Item opjosdhipote and the Definite Article. In C. Halpert, J. Hartman & D. Hill (Eds.), *Proceedings of the 2007 Workshop in Greek Syntax and Semantics at MIT*, MIT Working Papers in Linguistics 57 (pp. 123-137). Cambridge: MA.

Lazaridou-Chatzigoga, D., Katsos, N., & Stockall, L. (2015). Genericity is easy? Formal and experimental perspectives. In Hansen, N. & Borg, E. (Eds.), *RATIO*, 28(4), Special Issue: Investigating Meaning: Experimental Approaches, 470-494.

Lazaridou-Chatzigoga, D. & Stockall, L. (2013). Genericity, exceptions and domain restriction: experimental evidence from comparison with universals. In E. Chemla, V. Homer, & G. Winterstein (Eds.), Proceedings of Sinn und Bedeutung 17 (pp. 325-343). École Normale Supérieure, Paris.

Lazaridou-Chatzigoga, D., Stockall, L., & Katsos, N. (2017). A new look at the 'Generic Overgeneralisation' effect, *Inquiry*. DOI: 10.1080/0020174X.2017.1285993.

Leslie, S. J. (2007). Generics and the structure of the mind. *Philosophical Perspectives*, *21*, 375-405.

Leslie, S. J. (2008). Generics: Cognition and acquisition. *The Philosophical Review*, *117(1)*, 1-49.

Leslie, S. J., Khemlani, S., & Glucksberg, S. (2011). All Ducks Lay Eggs: The Generic Overgeneralization Effect. *Journal of Memory and Language*, *65(1)*, 15-31.

Longobardi, G. (2001). How comparative is semantics? A unified parametric theory of bare nouns and proper names. *Natural Language Semantics*, *9(4)*, 335-369.

Mannheim, B., Gelman, S. A., Escalante, C., Huayhua, M., & Puma, R. (2011). A Developmental Analysis of Generic Nouns in Southern Peruvian Quechua. *Language Learning and Development*, *7(1)*, 1-23.

Mari, A., Beyssade, C., & del Prete, F. (2013). *Genericity*. Oxford: Oxford University Press.

Marmaridou-Protopapa, S. (1984). The study of reference, attribution and genericness in the

context of English and their grammaticalization in M. Greek noun phrases. Unpublished Ph.D. thesis. Darwin College. Cambridge.

Martí, L. (2003). *Contextual variables*. PhD dissertation, University of Connecticut.

Matthewson, L. (2001). Quantification and the nature of crosslinguistic variation. *Natural Language Semantics*, *9*, 145-179.

Meyer, M., Gelman, S. A., & Stilwell, S.M. (2011). Generics are a cognitive default: Evidence from sentence processing. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 913-918). Boston, MA: Cognitive Science Society.

Nickel, B. (2008). Generics and the Ways of Normality. *Linguistics and Philosophy, 31(6)*, 629–48.

Partee, B. (1995). Quantificational structures and compositionality. In E. Bach, E. Jelinek, A. Kratzer & B. Partee (Eds.). *Quantification in Natural Languages* (pp. 541-601). Dordrecht: Kluwer.

Pelletier, F. & Asher, N. (1997). Generics and defaults. In J. van Benthem & A. ter Meulen (Eds.) *Handbook of Logic and Language* (pp. 1125–1179). Cambridge, MA: MIT Press.

Pesetsky, D. (1987). Wh-in-Situ: Movement and Unselective Binding. In E. J. Reuland & A. G. B. ter Meulen (Eds.), *The Representation of (In)definiteness* (pp. 98-129). Cambridge, MA: MIT Press.

Prasada, S., Khemlani, S., Leslie, S. J., & Glucksberg, S. (2013). Conceptual distinctions amongst generics. *Cognition, 126*, 405-422.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Recanati, F. (1996). Domains of discourse. *Linguistics and Philosophy*, *19*, 445-475.

Roussou, A. & Tsimpli, I.-M. (1994). On the interaction of case and definiteness in Modern Greek. In I. Philippaki-Warburton, K. Nicolaidis & M. Sifianou (Eds.), *Themes in Greek Linguistics* (pp. 69-76). Amsterdam: John Benjamins.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic Memory: A featural model for semantic decisions. *Psychological Review*, 81, 214-241.

Stanley, J. (2002). Nominal restriction. In G. Preyer & G. Peters, *Logical Form and Language* (pp. 365-388). Oxford University Press.

Stanley, J., & Szabó, Z.G. (2000). On Quantifier Domain Restriction. *Mind and Language*, *15 (2-3)*, 219-261.

Sterken, R. (2015). Generics in Context. *Philosopher's Imprint*, *15(21)*, 1-30.

Tardif, T., Gelman, S.A., Fu, X., & Zhu, L. (2012). Acquisition of generic noun phrases in Chinese: learning about lions without an '-s'. *Journal of Child Language, 39*, 130-161.

Tsamadou-Jacoberger, I. (2006). A propos de ΟΛΟΣ 'TOUT' en grec moderne. In G. Kleiber, C. Schnedecker, & A. Theissen (Eds), *La relation partie-tout, Bibliothèque de l'Information Grammaticale*, 59 (pp. 227-241), Louvain-Paris, Editions Peters.

Tsili, M. (2001). The quantificational phrases o (kathe) NP, (o) kathenas > universality and distributivity [In Greek]. *Studies in Greek Linguistics*, *21* (pp. 783-794). Thessaloniki: Kyriakidis.

Vendler, Z. (1962). Each and every, any and all. In *Mind*, *71*, 145-160. Reprinted in Z. Vendler (1967) *Linguistics in Philosophy* (pp. 70-96). Ithaca: NY: Cornell University Press.

Vergnaud, J.-R., & Zubizarreta, M.L. (1992). The definite determiner and the inalienable constructions in French and English. *Linguistic Inquiry*, *23*, 595-652.

Appendix A
Materials for Experiments 1 and 2. Contexts did not vary across determiner types. The statements are given here in the generic form. Most of the statements are a subset of the items used in Leslie et al. (2011). For the test items, each participant saw only one type of context for a given statement. For the control items, there was only one type of context, which did not vary across conditions.

Experiment 1: English
Test items
Majority characteristic statements

1. **Background**:
   a. **neutral**
   At Stratford's Farm you can enjoy watching thirty of the world's most spectacular butterflies feeding from the wildflower meadow.
   b. **contradictory**
   At Stratford's Farm you can learn by watching researchers taking care of thirty mutant butterflies that never develop a means of flying.
   c. **supportive**
   At Stratford's Farm you can enjoy watching thirty of the world's most spectacular butterflies flying all around in a tropical environment.
   Statement:
   Butterflies have wings.

2. **Background**:
   a. **neutral**
   Cat Rescue takes in and cares for unwanted and abandoned animals. They now look after four cats that can be adopted by responsible owners.
   b. **contradictory**
   Cat Rescue takes in and cares for unwanted animals. They now look after four cats who have been in a fight and lost their bristles around the mouth.
   c. **supportive**
   Cat Rescue takes in and cares for unwanted and abandoned animals. They now look after four cats, whose long bristles make them really adorable.
   Statement:
   Cats have whiskers.

3. **Background**:
   a. **neutral**
   Chester Zoo is home to several African animals, among them five cheetahs. They were shipped from Kenya after detailed health checks.
   b. **contradictory**
   Chester Zoo is home to several African animals, among them five cheetahs. They have hurt their legs badly and have been in recovery for months.
   c. **supportive**
   Chester Zoo is home to several African animals, among them five cheetahs. They were able to go quicker than two Olympic sprinters in a race.
   Statement:
   Cheetahs run fast.

4. **Background**:
   a. **neutral**
   Row Farm is a family farm, where children have the chance to see, meet and take photos of farm animals, including twenty cows.
   b. **contradictory**
   Row Farm is a family farm, where children have the chance to see the twenty cows which are only fed with corn and other grains.
   c. **supportive**
   Row Farm is a family farm, where children have the chance to see the twenty cows that roam freely and enjoy the green pastureland.
   Statement:
   Cows eat grass.

5. **Background**:
   a. **neutral**
   Shakespeare Tower at the Barbican Estate has many older residents who keep pets. Ten dogs have been acquired recently in the building.
   b. **contradictory**
   Shakespeare Tower at the Barbican Estate has many older residents who keep pets. Ten dogs have recently gone through cosmetic docking.
   c. **supportive**
   Shakespeare Tower at the Barbican Estate has many older residents who keep pets. Ten dogs are taken care of, happily moving their flexible appendage.
   Statement:
   Dogs have tails.

6. **Background**:
   a. **neutral**
   The Riding Club and Livery Stables in the heart of Wimbledon owns five horses, which are well trained in offering enjoyable rides.
   b. **contradictory**
   The Riding Club and Livery Stables in the heart of Wimbledon owns five horses, which had accidents and have had hind limbs amputated.
   c. **supportive**
   The Riding Club and Livery Stables in the heart of Wimbledon owns five horses, which are ready to offer visitors rides on their healthy limbs.
   Statement:
   Horses have four legs.

7. **Background**:
   a. **neutral**
   At Trealy Farm visitors can observe and enjoy fifteen rabbits they take care of, which are lucky to have found a safe urban home.
   b. **contradictory**
   At Trealy Farm visitors can observe fifteen disabled rabbits they take care of, which are lucky to receive treatment for their paralysis.
   c. **supportive**
   At Trealy Farm visitors can observe and enjoy fifteen rabbits, which have a special way of jumping due to their long and strong back legs.
   Statement:

Rabbits hop.

8. **Background**:
   a. **neutral**
   Dudley Zoo has three leopards, which are popular with visitors, who are fascinated with their sleek, powerful bodies and their swimming skills.
   b. **contradictory**
   Dudley Zoo has three leopards, which are popular with visitors, who are fascinated because they were born without any marks on their bodies.
   c. **supportive**
   Dudley Zoo has three leopards, which are popular with visitors, who are fascinated with their full black markings and their rosette pattern.
   Statement:
   Leopards have spots.

9. **Background**:
   a. **neutral**
   Residents enjoy the birds that can be found in Central Park. One day, ten sparrows were spotted there, chirping and flying around happily.
   b. **contradictory**
   Residents enjoy the birds that can be found in Central Park. One day, ten sparrows got trapped in wires resulting in total separation of their bills.
   c. **supportive**
   Residents enjoy the birds that can be found in Central Park. One day, ten sparrows were spotted there, with a bright yellow lower half of the bill.
   Statement:
   Sparrows have beaks.

10. **Background**:
    a. **neutral**
    Linton Zoo is home to three tigers, Tibor, Baginda and Kaytlin, whose playful games visitors love to watch and photograph.
    b. **contradictory**
    Linton Zoo is home to three tigers, Tibor, Baginda and Kaytlin, whose fur is all white due to a recessive gene that controls coat color.
    c. **supportive**
    Linton Zoo is home to three tigers, Tibor, Baginda and Kaytlin, whose black and orange coats visitors love to photograph.
    Statement:
    Tigers have stripes.

11. **Background**:
    a. **neutral**
    The California Sanctuary takes care of five elephants which enjoy year-round access to their beautiful and expansive natural habitat.
    b. **contradictory**
    The California Sanctuary takes care of five elephants which have had their elongated nasal part amputated due to crocodile attacks.
    c. **supportive**

46

The California Sanctuary takes care of five elephants which push down trees and pick up straws with their unique elongated nasal part.
Statement:
Elephants have trunks.

12. **Background**:
a. **neutral**
In Battersea Animal Shelter they have four elderly squirrels, which are so tame that they might come and sit on your lap.
b. **contradictory**
In Battersea Animal Shelter they have four elderly toothless squirrels, which do not eat anything with a hard or soft shell.
c. **supportive**
In Battersea Animal Shelter they have four adult squirrels, which run up and down and hold their beloved acorns in their tiny paws.
Statement:
Squirrels eat nuts.

Control items
Definitional statements

1. **Background**:
The New Bennett School is used to film a story on retro pets. To prepare for the visit, they have two farms up and running with 20 ants.
Statement:
Ants are insects.

2. **Background**:
A Greek journalist has taken an interest and compiled a list of Ireland's 50 most eligible bachelors, which will be published online.
Statement:
Bachelors are unmarried.

3. **Background**:
Pets are a wonderful addition to households with children, since they are fun to have around. In a central block of flats, you can find up to ten mice.
Statement:
Mice are animals.

4. **Background**:
The Safari Park hosts many great apes, among them five popular gorillas, who all peel radishes with their teeth before eating them.
Statement:
Gorillas are mammals.

5. **Background**:
The tourism board produced a colored leaflet with details about the fauna and flora of Greenwich Park, which now includes 30 elms.
Statement:
Elms are trees.

6. **Background**:
Outside the city, it is possible to walk in the woods and explore the rich habitat. One walker spotted 50 mushrooms after a rainy day.
Statement:
Mushrooms are fungi.

7. **Background**:
The RSPCA rehoming process makes sure pets and owners are well matched. More than 30 candidates were considered to adopt ten poodles.
Statement:
Poodles are dogs.

8. **Background**:
Outside the window of a family in Bristol, three robins have built their nest. This is the second year they have used the same spot.
Statement:
Robins are birds.

9. **Background**:
Teachers in Kent engage their pupils in maths lessons using fun and interactive activities to create triangles with threads.
Statement:
Triangles have three sides.

10. **Background**:
Queen Mary's Garden is right in the middle of the city and visitors can enjoy one of the best rose collections in the country for free.
Statement:
Roses are flowers.

11. **Background**:
In the Animal Corner you can learn amazing facts about the twenty cobras they host, which can lift a third of their body off the ground.
Statement:
Cobras are snakes.

12. **Background**:
Soaring temperatures in England have allowed the sea off the Essex coast to warm. A local angler caught fifty anchovies for the first time.
Statement:
Anchovies are fish.

False generalisation statements

13. **Background**:
A visit to the Natural History Museum can motivate children and prepare them to write essays on animals based on what they learned.
Statement:
Animals are reptiles.

14. **Background**:
    The 2012 Summer Olympic Games took place in London, UK from 27 July to 12 August, with an estimated 2000 athletes participating.
    Statement:
    Athletes are students.

15. **Background**:
    Paynes Southdown Bee Farms Limited operate hundreds of hives all along the South Downs area. They are estimated to have over 3000 bees.
    Statement:
    Bees are worker bees.

16. **Background**:
    Blackwell's was the first bookstore in the UK that allowed its customers to purchase books online, with access to 150,000 titles.
    Statement:
    Books are paperbacks.

17. **Background**:
    The Pavillion overlooks the Boating Lake, where you can rent a boat and play with the water. There are currently four swans in the lake.
    Statement:
    Swans are female.

18. **Background**:
    Epping Forest offers numerous activities and events across the year. There are currently ten girls performing at the Open Air Theater.
    Statement:
    Girls have curly hair.

19. **Background**:
    Wildlife officials noticed an increase in calls lately. Ten bears foraging for food were seen in neighborhoods across southern Maine.
    Statement:
    Bears have white fur.

20. **Background**:
    Cape Royds is a rocky promontory overlaid with dirty ice. Over fifty penguins were spotted there, which struggle to cope with global warming.
    Statement:
    Penguins are male.

21. **Background**:
    A new idea is implemented in the Millenium Park that brings together the edible and incredible: veggies, herbs and flowers in one garden.
    Statement:
    Flowers are yellow.

22. **Background**:

Queens' Park's revitalization plan involved extensive planting and the improvement of the Park entrance. 3000 trees were planted.
Statement:
Trees are deciduous.

23. **Background**:
On Bow Street there are currently twenty cars parked. Many disputes arise over the failure to observe parking 'etiquette'.
Statement:
Cars are black.

24. **Background**:
On Old Kent Road there are thirty houses, which were all built during the last decade. People from ten different nationalities live there.
Statement:
Houses are mansions.

Experiment 2: Greek
Test items
Majority characteristic statements

1. **Background**:
a. **neutral**
Στην φάρμα Κουκάκη μπορείτε να δείτε τριάντα από τις πιο εντυπωσιακές πεταλούδες του κόσμου, που τρέφονται από το λιβάδι με τα αγριολούλουδα.
[English translation: At Koukaki Farm you can see thirty of the world's most spectacular butterflies feeding from the wildflower meadow]
b. **contradictory**
Στην φάρμα Κουκάκη μπορείτε να παρακολουθήσετε ερευνητές να φροντίζουν τριάντα μεταλλαγμένες πεταλούδες, που δεν αναπτύσσουν ποτέ την δυνατότητα να πετάξουν.
[English translation: At Koukaki Farm you can watch researchers taking care of thirty mutant butterflies that never develop a means of flying.]
c. **supportive**
Στην φάρμα Κουκάκη μπορείτε να δείτε και να απολαύσετε τριάντα από τις πιο εντυπωσιακές πεταλούδες του κόσμου να πετάνε γύρω γύρω σε ένα τροπικό περιβάλλον.
[English translation: At Koukaki Farm you can see and enjoy thirty of the world's most spectacular butterflies flying all around in a tropical environment.]
Statement:
Οι πεταλούδες έχουν φτερά.
[English translation: Butterflies have wings.]

2. **Background**:
a. **neutral**
Το σωματείο Σώζω φροντίζει ανεπιθύμητα και εγκαταλελειμμένα ζώα. Τώρα φιλοξενούν τέσσερις γάτες που μπορούν να υιοθετηθούν από υπεύθυνους ιδιοκτήτες.
[English translation: The club 'I save' cares for unwanted and abandoned animals. They now host four cats that can be adopted by responsible owners.]

b. **contradictory**

Το σωματείο Σώζω φροντίζει ανεπιθύμητα ζώα. Τώρα φιλοξενούν τέσσερις γάτες που πήραν μέρος σε έναν καυγά και έχασαν τις τρίχες γύρω από το στόμα τους.

[English translation: The club 'I save' cares for unwanted and abandoned animals. They now host four cats who have been in a fight and lost their bristles around the mouth.]

c. **supportive**

Το σωματείο Σώζω φροντίζει ανεπιθύμητα ζώα. Τώρα φιλοξενούν τέσσερις γάτες, που οι καμπυλωτές τρίχες στο πάνω χείλος τους τις κάνουν αξιαγάπητες.

[English translation: The club 'I save' cares for unwanted and abandoned animals. They now host four cats, whose curvy bristles around their upper lip make them really adorable.]

Statement:

Οι γάτες έχουν μουστάκια.

[English translation: Cats have whiskers.]

3. **Background**:

a. **neutral**

Ο ζωολογικός κήπος του Τσέστερ έχει πολλά ζώα από την Αφρική, ανάμεσά τους και πέντε τσίτα. Τα έστειλαν από την Κένυα μετά από ενδελεχή έλεγχο για την υγεία τους.

[English translation: Chester Zoo is home to several African animals, among them five cheetahs. They were shipped from Kenya after detailed health checks.]

b. **contradictory**

Ο ζωολογικός κήπος του Τσέστερ έχει πολλά ζώα από την Αφρική, ανάμεσά τους και πέντε τσίτα που χτύπησαν άσχημα τα πόδια τους και είναι σε ανάρρωση εδώ και μήνες.

[English translation: Chester Zoo is home to several African animals, among them five cheetahs. They have hurt their legs badly and have been in recovery for months.]

c. **supportive**

Ο ζωολογικός κήπος του Τσέστερ έχει πολλά ζώα από την Αφρική, ανάμεσά τους και πέντε τσίτα που σε έναν αγώνα κατάφεραν να πάνε πιο γρήγορα από δύο Ολυμπιακούς αθλητές.

[English translation: Chester Zoo is home to several African animals, among them five cheetahs. They were able to go quicker than two Olympic sprinters in a race.]

Statement:

Τα τσίτα τρέχουν γρήγορα.

[English translation: Cheetahs run fast.]

4. **Background**:

a. **neutral**

Η φάρμα Σκαφιδά είναι μία οικογενειακή φάρμα, όπου τα παιδιά έχουν την ευκαιρία να δουν και να φωτογραφήσουν ζώα της φάρμας, συμπεριλαμβανομένων είκοσι αγελάδων.

[English translation: Skafida Farm is a family farm, where children have the chance to see and take photos of farm animals, including twenty cows.]

b. **contradictory**

Η φάρμα Σκαφιδά είναι μία οικογενειακή φάρμα, όπου τα παιδιά μπορούν να δουν τις είκοσι αγελάδες που έχουν, οι οποίες τρώνε μόνο καλαμπόκι και άλλα σιτηρά.

[English translation: Skafida Farm is a family farm, where children have the chance to see the twenty cows which are only fed with corn and other grains.]
c. **supportive**
Η φάρμα Σκαφιδά είναι μία οικογενειακή φάρμα, όπου τα παιδιά μπορούν να δουν τις είκοσι αγελάδες που έχουν, οι οποίες βόσκουν ελεύθερα και απολαμβάνουν το πράσινο βοσκοτόπι.
[English translation: Skafida Farm is a family farm, where children have the chance to see the twenty cows that they have, which roam freely and enjoy the green pastureland.]
Statement:
Οι αγελάδες τρώνε γρασίδι.
[English translation: Cows eat grass.]

5. **Background**:
   a. **neutral**
   Το κεντρικό κτίριο του συμπλέγματος Κοκκινάκη έχει πολλούς ηλικιωμένους κατοίκους που έχουν κατοικίδια. Πρόσφατα, προστέθηκαν δέκα σκυλιά στο κτίριο.
   [English translation: The main building of the Kokkinaki estate has many older residents who keep pets. Ten dogs have been added recently to the building.]
   b. **contradictory**
   Το κεντρικό κτίριο του συμπλέγματος Κοκκινάκη έχει πολλούς ηλικιωμένους κατοίκους που έχουν κατοικίδια. Δέκα σκυλιά πρόσφατα υπεβλήθησαν σε καλλωπιστικό ακρωτηριασμό.
   [English translation: The main building of the Kokkinaki estate has many older residents who keep pets. Ten dogs have recently gone through cosmetic docking.]
   c. **supportive**
   Το κεντρικό κτίριο του συμπλέγματος Κοκκινάκη έχει πολλούς ηλικιωμένους κατοίκους που έχουν κατοικίδια. Οι κάτοικοι φροντίζουν δέκα σκυλιά που κουνούν χαρούμενα την ευλύγιστη απόφυσή τους.
   [English translation: The main building of the Kokkinaki estate has many older residents who keep pets. The residents take care of ten dogs who move happily their flexible appendage.]
   Statement:
   Οι σκύλοι έχουν ουρά.
   [English translation: Dogs have tails.]

6. **Background**:
   a. **neutral**
   Ο Ιππικός Όμιλος και Σταύλος Σαραντικού στην καρδιά του Ηρακλείου φιλοξενεί πέντε άλογα, που είναι καλά εκπαιδευμένα και προσφέρουν απολαυστικές βόλτες.
   [English translation: The Riding Club and Stables of Sarantikos in the heart of Herakleion hosts five horses, which are well trained and offer enjoyable rides.]
   b. **contradictory**
   Ο Ιππικός Όμιλος και Σταύλος Σαραντικού στην καρδιά του Ηρακλείου έχει πέντε άλογα, που είχαν ατυχήματα και έχασαν τα οπίσθια άκρα τους.
   [English translation: The Riding Club and Stables of Sarantikos in the heart of Herakleion owns five horses, which had accidents and have had hind limbs amputated.]
   c. **supportive**

52

Ο Ιππικός Όμιλος και Σταύλος Σαραντικού στην καρδιά του Ηρακλείου έχει πέντε άλογα, που είναι έτοιμα να προσφέρουν στους επισκέπτες βόλτες πάνω στα υγιή άκρα τους.
[English translation: The Riding Club and Stables of Sarantikos in the heart of Herakleion owns five horses, which are ready to offer visitors rides on their healthy limbs.]
Statement:
Τα άλογα έχουν τέσσερα πόδια.
[English translation: Horses have four legs.]

7. **Background**:
a. **neutral**
Στην φάρμα Τρικάλων μπορεί κανείς να παρακολουθήσει τα δεκαπέντε κουνέλια που φροντίζουν, τα οποία είναι τυχερά που βρήκαν ένα ασφαλές σπίτι στην πόλη.
[English translation: At Trikala Farm you can observe and enjoy the fifteen rabbits they take care of, which are lucky to have found a safe urban home.]
b. **contradictory**
Στην φάρμα Τρικάλων μπορεί κανείς να παρακολουθήσει τα δεκαπέντε ανάπηρα κουνέλια, που έχουν την τύχη να θεραπεύονται για την παράλυσή τους.
[English translation: At Trikala Farm you can observe the fifteen disabled rabbits, which are lucky to receive treatment for their paralysis.]
c. **supportive**
Στην φάρμα Τρικάλων μπορεί κανείς να παρακολουθήσει τα δεκαπέντε κουνέλια που κινούνται με έναν ιδιαίτερο τρόπο εξαιτίας του μήκους και της δύναμης των ποδιών τους.
[English translation: At Trikala Farm you can observe the fifteen rabbits which have a special way of jumping due to their long and strong back legs.]
Statement:
Τα κουνέλια χοροπηδάνε.
[English translation: Rabbits hop.]

8. **Background**:
a. **neutral**
Ο ζωολογικός κήπος του Λόβετς έχει τρεις λεοπαρδάλεις που είναι δημοφιλείς μεταξύ των επισκεπτών, οι οποίοι εντυπωσιάζονται από τα λεία σώματά τους και τις κολυμβητικές τους ικανότητες.
[English translation: Lovech Zoo has three leopards, which are popular with visitors, who are fascinated with their sleek, powerful bodies and their swimming skills.]
b. **contradictory**
Ο ζωολογικός κήπος του Λόβετς έχει τρεις λεοπαρδάλεις που είναι δημοφιλείς μεταξύ των επισκεπτών, οι οποίοι εντυπωσιάζονται γιατί γεννήθηκαν χωρίς καθόλου σημάδια στο σώμα τους.
[English translation: Lovech Zoo has three leopards, which are popular with visitors, who are fascinated because they were born without any marks on their bodies.]
c. **supportive**
Ο ζωολογικός κήπος του Λόβετς έχει τρεις λεοπαρδάλεις που είναι δημοφιλείς μεταξύ των επισκεπτών, οι οποίοι εντυπωσιάζονται με τα ολόμαυρα σημάδια στο σώμα τους.
[English translation: Lovech Zoo has three leopards, which are popular with visitors, who are fascinated with their full black markings on their bodies.]

Statement:
Οι λεοπαρδάλεις έχουν βούλες.
[English translation: Leopards have spots.]

9. **Background**:
   a. **neutral**
   Οι κάτοικοι της περιοχής απολαμβάνουν τα πουλιά που μπορεί να δει κανείς στο Σέιχ Σου. Μια μέρα εντόπισαν δέκα σπουργίτια να κελαηδούν και να πετούν τριγύρω.
   [English translation: The area's residents enjoy the birds that one can see in Seih Sou. One day, ten sparrows were spotted there, chirping and flying around happily.]
   b. **contradictory**
   Οι κάτοικοι της περιοχής απολαμβάνουν τα πουλιά που μπορεί να δει κανείς στο Σέιχ Σου. Μια μέρα εντόπισαν δέκα παγιδευμένα σπουργίτια που το στόμα τους είχε αποκολληθεί εντελώς.
   [English translation: The area's residents enjoy the birds that one can see in Seih Sou. One day, they spotted ten trapped sparrows whose mouth had been separated completely.]
   c. **supportive**
   Οι κάτοικοι της περιοχής απολαμβάνουν τα πουλιά που μπορεί να δει κανείς στο Σέιχ Σου. Μια μέρα εντόπισαν δέκα σπουργίτια με έντονο κίτρινο στόμα.
   [English translation: The area's residents enjoy the birds that one can see in Seih Sou. One day, they spotted ten sparrows with a bright yellow mouth.]
   Statement:
   Τα σπουργίτια έχουν ράμφος.
   [English translation: Sparrows have beaks.]

10. **Background**:
    a. **neutral**
    Ο ζωολογικός κήπος της Κριμαίας φιλοξενεί τρεις τίγρεις, τον Τίμπορ, την Μπαγκίντα και την Κάιτλιν, που ξεσηκώνουν τους επισκέπτες με τα αστεία παιχνίδια τους.
    [English translation: Crimea Zoo is home to three tigers, Tibor, Baginda and Kaytlin, whose funny games excite visitors.]
    b. **contradictory**
    Ο ζωολογικός κήπος της Κριμαίας φιλοξενεί τρεις τίγρεις, τον Τίμπορ, την Μπαγκίντα και την Κάιτλιν, που το τρίχωμά τους είναι λευκό εξαιτίας ενός γονιδίου που ελέγχει το χρώμα της γούνας.
    [English translation: Crimea Zoo is home to three tigers, Tibor, Baginda and Kaytlin, whose fur is white due to a recessive gene that controls coat color.]
    c. **supportive**
    Ο ζωολογικός κήπος της Κριμαίας φιλοξενεί τρεις τίγρεις, τον Τίμπορ, την Μπαγκίντα και την Κάιτλιν, των οποίων το πορτοκαλόμαυρο τρίχωμα θαυμάζουν οι επισκέπτες.
    [English translation: Crimea Zoo is home to three tigers, Tibor, Baginda and Kaytlin, whose black and orange coats visitors admire.]
    Statement:
    Οι τίγρεις έχουν ρίγες.
    [English translation: Tigers have stripes.]

11. **Background**:

a. **neutral**

Το καταφύγιο της Καλιφόρνια φροντίζει πέντε ελέφαντες που έχουν πρόσβαση στο όμορφο και μεγάλο φυσικό περιβάλλον όλο το χρόνο.

[English translation: The California Sanctuary takes care of five elephants which enjoy year-round access to their beautiful and expansive natural habitat.]

b. **contradictory**

Το καταφύγιο της Καλιφόρνια φροντίζει πέντε ελέφαντες που έχασαν το επίμηκες ρινικό μέλος τους εξαιτίας μίας επίθεσης κροκοδείλων.

[English translation: The California Sanctuary takes care of five elephants, which have lost their elongated nasal part due to a crocodile attack.]

c. **supportive**

Το καταφύγιο της Καλιφόρνια φροντίζει πέντε ελέφαντες που ρίχνουν κάτω δέντρα και σηκώνουν άχυρα με το επίμηκες ρινικό μέλος τους.

[English translation: The California Sanctuary takes care of five elephants which push down trees and pick up straws with their unique elongated nasal part.]

Statement:

Οι ελέφαντες έχουν προβοσκίδα.

[English translation: Elephants have trunks.]

12. **Background**:

a. **neutral**

Στο Πάρκο Αγρινίου υπάρχουν τέσσερις ηλικιωμένοι σκίουροι, οι οποίοι είναι τόσο εξημερωμένοι που μπορεί να έρθουν και να καθίσουν πάνω σου.

[English translation: In Agrinio Park there are four elderly squirrels, which are so tame that they might come and sit on your lap.]

b. **contradictory**

Στο Πάρκο Αγρινίου υπάρχουν τέσσερις ηλικιωμένοι σκίουροι χωρίς δόντια, που δεν τρώνε τίποτα που να έχει σκληρό ή μαλακό κέλυφος.

[English translation: In Agrinio Park there are four elderly squirrels, which do not eat anything with a hard or soft shell.]

c. **supportive**

Στο Πάρκο Αγρινίου υπάρχουν τέσσερις μεγάλοι σκίουροι που τρέχουν πάνω κάτω και κρατάνε τα αγαπημένα τους βελανίδια στα μικρά πόδια τους.

[English translation: In Agrinio Park there are four elderly squirrels, which run up and down and hold their beloved acorns in their tiny paws.]

Statement:

Οι σκίουροι τρώνε καρπούς.

[English translation: Squirrels eat nuts.]

Control items
Definitional statements

1. **Background**:

Το Πρότυπο Σχολείο χρησιμοποιείται για να γυριστεί ένα επεισόδιο για ρετρό κατοικίδια. Για να προετοιμαστούν για το γύρισμα, έφτιαξαν δύο φωλιές μυρμηγκιών.

[English translation: The Model School is used to film an episode on retro pets. To prepare for the visit, they built two farms with ants.]

Statement:

Τα μυρμήγκια είναι έντομα.

[English translation: Ants are insects.]

2. **Background**:
Ένας Έλληνας δημοσιογράφος σκέφτηκε να συλλέξει μία λίστα με τους 50 πιο ελκυστικούς εργένηδες της Ιρλανδίας, που θα δημοσιευτεί στο διαδίκτυο.
[English translation: A Greek journalist thought of compiling a list of Ireland's 50 most eligible bachelors, which will be published on the internet.]
Statement:
Οι εργένηδες είναι ανύπαντροι.
[English translation: Bachelors are unmarried.]

3. **Background**:
Τα κατοικίδια είναι μία υπέροχη προσθήκη σε σπιτικά με παιδιά. Σε μία κεντρική πολυκατοικία, μπορείς να βρεις μέχρι και δέκα ποντίκια, που είναι πολύ ευχάριστη παρέα.
[English translation: Pets are a wonderful addition to households with children. In a central block of flats, you can find up to ten mice, who keep a nice company.]
Statement:
Τα ποντίκια είναι ζώα.
[English translation: Mice are animals.]

4. **Background**:
Το Πάρκο Σαφάρι φιλοξενεί πολλούς μεγάλους πιθήκους, ανάμεσά τους πέντε γορίλες, που ξεφλουδίζουν τα ραπανάκια με τα δόντια τους πριν τα φάνε.
[English translation: The Safari Park hosts many great apes, among them five popular gorillas, who all peel radishes with their teeth before eating them.]
Statement:
Οι γορίλες είναι θηλαστικά.
[English translation: Gorillas are mammals.]

5. **Background**:
Το τουριστικό γραφείο έφτιαξε ένα έγχρωμο φυλλάδιο με λεπτομέρειες σχετικά με την χλωρίδα και την πανίδα του Εθνικού Κήπου που περιλαμβάνει τριάντα πλατάνια.
[English translation: The tourism board produced a colored leaflet with details about the fauna and flora of the National Park, which now includes 30 plane trees.]
Statement:
Τα πλατάνια είναι δέντρα.
[Engligh translation: Plane trees are tress.]

6. **Background**:
Έξω από την πόλη, μπορεί κανείς να περπατήσει στο δάσος και να εξερευνήσει το πλούσιο οικοσύστημα. Ένας περαστικός εντόπισε πενήντα μανιτάρια.
[English translation: Outside the city, it is possible to walk in the woods and explore the rich habitat. One walker spotted 50 mushrooms after a rainy day.]
Statement:
Τα μανιτάρια είναι μύκητες.
[English translation: Mushrooms are fungi.]

7. **Background**:

Το πρόγραμμα της Φιλοζωικής φροντίζει να βρίσκει κατάλληλους ιδιοκτήτες για τα κατοικίδια. Περισσότεροι από 30 υποψήφιοι εξετάστηκαν για να υιοθετήσουν μόλις δέκα κανίς.

[English translation: The Animal Welfare Society's program makes sure to find adequate owners for the pets. More than 30 candidates were considered to adopt ten poodles.]

Statement:

Τα κανίς είναι σκυλιά.

[English translation: Poodles are dogs.]

8. **Background**:

Έξω από το παράθυρο μιας οικογένειας στην Δράμα, τρία χελιδόνια έφτιαξαν την φωλιά τους. Είναι η δεύτερη χρονιά που χρησιμοποιούν το ίδιο σημείο.

[English translation: Outside the window of a family in Drama, three swallows have built their nest. This is the second year they have used the same spot.]

Statement:

Τα χελιδόνια είναι πουλιά.

[English translation: Swallows are birds.]

9. **Background**:

Ο ιππικός Όμιλος Κυκλάδων έχει στο δυναμικό του σαράντα άλογα και προσφέρει μαθήματα σε αρχάριους και προχωρημένους ιππείς.

[English translation: The Cyclades Riding club has forty horses in its force and offers lessons for beginner and advanced riders.]

Statement:

Οι φοράδες είναι άλογα.

[English translation: Female horses are horses.]

10. **Background**:

Το εκτροφείο "ο Παράδεισος" εξειδικεύεται σε ζώα ταϊσμένα στο χέρι. Ανάμεσά τους υπάρχουν τριάντα παπαγάλοι από τροπικές και υποτροπικές χώρες.

[English translation: The breeding group 'The Paradise' specialises in animals fed by hand. Among these, there are thirty parrots from tropical and subtropical countries.]

Statement:

Οι παπαγάλοι είναι πτηνά.

[English translation: Parrots are birds.]

11. **Background**:

Στην Γωνιά του Ζώου μπορείς να μάθεις απίστευτα στοιχεία για τις είκοσι κόμπρες που φιλοξενούν, που μπορούν να σηκώσουν το ένα τρίτο του σώματός τους πάνω από το έδαφος.

[English translation: In the Animal Corner you can learn amazing facts about the twenty cobras they host, which can lift a third of their body off the ground.]

Statement:

Οι κόμπρες είναι φίδια.

[English translation: Cobras are snakes.]

12. **Background**:

Οι υψηλές θερμοκρασίες στην Αγγλία επέτρεψαν στην θάλασσα κοντά στο Έσσεξ να ζεσταθεί. Ένας ντόπιος ψαράς έπιασε πενήντα αντζούγιες για πρώτη φορά.

[English translation: Soaring temperatures in England have allowed the sea off the Essex coast to warm. A local angler caught fifty anchovies for the first time.]
Statement:
Οι αντσούγιες είναι ψάρια.
[English translation: Anchovies are fish.]

False generalisation statements

13. **Background**:
Μία επίσκεψη στο Μουσείο Φυσικής Ιστορίας μπορεί να ξεσηκώσει τα παιδιά και να τα προετοιμάσει να γράψουν εργασίες για όλα τα είδη ζώων που βλέπουν εκεί.
[English translation: A visit to the Natural History Museum can motivate children and prepare them to write essays on all the kinds of animals they see there].
Statement:
Τα ζώα είναι ερπετά.
[English translation: Animals are reptiles.]

14. **Background**:
Στους Ολυμπιακούς Αγώνες του Λονδίνου το 2012 πήραν μέρος 2000 αθλητές όλων των ηλικιών σε πάνω από σαράντα διαφορετικά αθλήματα.
[English translation: In the 2012 Summer Olympic Games 2000 athletes of all ages took part in more than forty different sports.]
Statement:
Οι αθλητές είναι φοιτητές.
[English translation: Athletes are students.]

15. **Background**:
Στην έκθεση νέων σχεδιαστών μπορεί κανείς να βρει κάθε λογής τραπέζια, με ιδιαίτερο στυλ και ανορθόδοξες διαστάσεις.
[English translation: In the new designers' exhibition one can find every kind of table, with unique style and unorthodox dimensions.]
Statement:
Τα τραπέζια είναι τετράγωνα.
[English translation: The tables are square.]

16. **Background**:
Στα βιβλιοπωλεία της Θεσσαλονίκης μπορεί κανείς να βρει βιβλία όλων των ειδών, από τις τελευταίες κυκλοφορίες μέχρι σπάνιες εκδόσεις.
[English translation: In Thessaloniki's bookstores you can find books of all kinds from the latest editions to rare editions.]
Statement:
Τα βιβλία έχουν μαλακό εξώφυλλο.
[English translation: Books are paperbacks.]

17. **Background**:
Στον Έβρο υπάρχουν πολλά είδη κύκνων που χαράζουν λευκές γραμμές στον ουρανό δημιουργώντας ένα μοναδικό θέαμα στην περιοχή.
[English translation: In Evros there are many kinds of swans that draw white lines in the sky creating a unique spectacle in the region.]
Statement:

Οι κύκνοι είναι θηλυκοί.
[English translation: Swans are female.]

18. **Background**:
Τα κορίτσια στην εφηβεία μπορεί να περάσουν ώρες μπροστά στον καθρέφτη δοκιμάζοντας διάφορα χτενίσματα και ρούχα ανάλογα με την περίσταση.
[English translation: Girls in adolescence may spend hours in front of a mirror trying out different hairstyles and clothes depending on the occasion.]
Statement:
Τα κορίτσια έχουν σγουρά μαλλιά.
[English translation: Girls have curly hair.]

19. **Background**:
Στην Ευρώπη μπορεί κανείς να δει τρία διαφορετικά είδη αρκούδας, πηγαίνοντας σε οργανωμένα σαφάρι που υπόσχονται μια περιπετειώδη διαδρομή.
[English translation: In Europe, one can see three different kinds of bear, going on an organised safari that promises an adventurous route.
Statement:
Οι αρκούδες έχουν λευκό τρίχωμα.
[English translation: Bears have white fur.]

20. **Background**:
Στο Κέιπ Ρόιντς μπορεί κανείς να επισκεφτεί το ερευνητικό κέντρο, όπου μελετούν δέκα διαφορετικά είδη πιγκουίνων που ζουν σε αυτό το βραχώδες ακρωτήριο.
[English translation: In Cape Royds one can visit the research centre, where they study ten different species of penguins, which live in this rocky promontory.]
Statement:
Οι πιγκουίνοι είναι αρσενικοί.
[English translation: Penguins are male.]

21. **Background**:
Το ετήσιο πάρτι των χρωμάτων και των αρωμάτων στήθηκε και φέτος στο Πάρκο Δρόσου, όπου η Ανθοκομική Έκθεση προσφέρει την καλύτερη βόλτα στην πόλη.
[English translation: The yearly party of colours and fragrances was set up this year in the Drosou Park, where the Flower Exhibition offers the best walk in town.]
Statement:
Τα λουλούδια είναι κίτρινα.
[English translation: Flowers are yellow.]

22. **Background**:
Μερικά δέντρα έχουν φύλλα που αλλάζουν χρώμα όταν μπαίνει το φθινόπωρο, ενώ άλλα είναι γεμάτα πράσινα φύλλα όλη την χρονιά.
[English translation: Some trees have leaves that change colour when the autum comes, while others are full of green leaves all year round.]
Statement:
Τα δέντρα είναι φυλλοβόλα.
[English translation: Trees are deciduous.]

23. **Background**:

Το γκαράζ στην οδό Ανάφης γεμίζει καθημερινά με αυτοκίνητα κάθε μάρκας και στυλ, καθώς επισκευάζουν κάθε εβδομάδα πάνω από τριάντα οχήματα.

[English translation: The garage in Anafi road fills up daily with cars of every brand and style, as they fix more than thirty vehicles every week.]

Statement:

Τα αυτοκίνητα είναι μαύρα.

[English translation: Cars are black.]

24. **Background**:

Για να χτίσει κανείς σπίτι σε οικόπεδο στο κέντρο της πόλης, πρέπει να έχει σχέδιο που να τηρεί τις πολεοδομικές διατάξεις της περιοχής.

[English translation: In order to build a house in a plot at the city centre, one has to have a plan that follows the urban planning regulations of the region.]

Statement:

Τα σπίτια είναι επαύλεις.

[English translation: Houses are mansions.]