Exploring Relations Between Ambiguities in Product Designs and Sentences

Eugine Szeto

Advised by Professor Shane Frederick

Submitted to the faculty of Cognitive Science in partial fulfillment of the requirements for the degree of Bachelor of Science

April 22nd, 2019

Author Contributions: Shane Frederick conceived of the original idea and provided thorough feedback throughout the development of the survey. Eugine Szeto helped implement the survey in addition to doing the background research, writing, and data analysis. Ryan Hauser helped with implementing the survey and data analysis.

Acknowledgements: Thank you to Mark Sheskin and the Yale Class of 2019 Cognitive Science majors for providing helpful feedback. Additional thank you to Hadas Kotek for providing guidance on language ambiguity.

Abstract

The world is often confusing because of ambiguous product designs and ambiguous sentences. Controls are not always naturally mapped to the object that is being controlled. Similarly, phrases in sentences can modify two different nouns. The goal of this paper is to explore the relationship between the following variables: how the ambiguities are perceived in products and in sentences (we measure participants' response time and confidence in their answers), how blame is assigned in the ambiguities, and psychological traits (IQ, CRT, and locus of control). We find that processing of ambiguity in product designs and sentences are likely to be related, and that IQ can be a mediating factor. Implications for product design and language relativity are discussed.

Introduction

Ambiguity is abundant in the world. It exists when a product may not have a proper label that shows how it can be used, and it exists when a sentence can have multiple meanings. When ambiguity exists, mistakes can occur. Depending on the context of that ambiguity, the consequences can vary.

A simple example is the design of a USB. How it is inserted into a port may be convoluted, and users may be frustrated, but they would probably move on from that momentary ordeal. However, if users are sufficiently frustrated, i.e. they do not find what they are looking for in a store or on a website, they may choose to take their business elsewhere (Krug, 2014; Underhill, 2009). Poor design can cause death or multiple deaths (Casey, 1998). In one case, three separate metal prongs – connected by wires to flat, circular objects taped to a four-year-old girl's chest – were accidentally plugged into an outlet on an intravenous pump instead of the heart monitor machine. Since the pump was connected to a wall socket, the Seattle nurse accidentally electrocuted and killed the young girl. In another case, grain containing methylmercury fungicide was imported into Iraq, and the packaging only contained a skull and bone as a warning against eating the grain. Thus, many did not understand the warning to stop themselves from eating the grain instead of planting it, which caused an estimated thousands of hospitalizations and hundreds of deaths.

In a history defining moment, a poor design – the butterfly ballot (see figure 1) – changed the outcome of the presidential election in 2000 (Wand, 2001). When the votes were tallied, the results indicated that an unusually large number of people voted for the third-party candidate, Buchanan. On the butterfly ballot, the Buchannan choice was second in the bubble choices, but that was only evident if one looks at both sides of the ballot. On the left side of the ballot, Gore appeared second. Thus, if one was looking to vote for Gore and was only looking at the left side of the ballot, one was likely to select the second bubble, inadvertently voting for Buchannan.

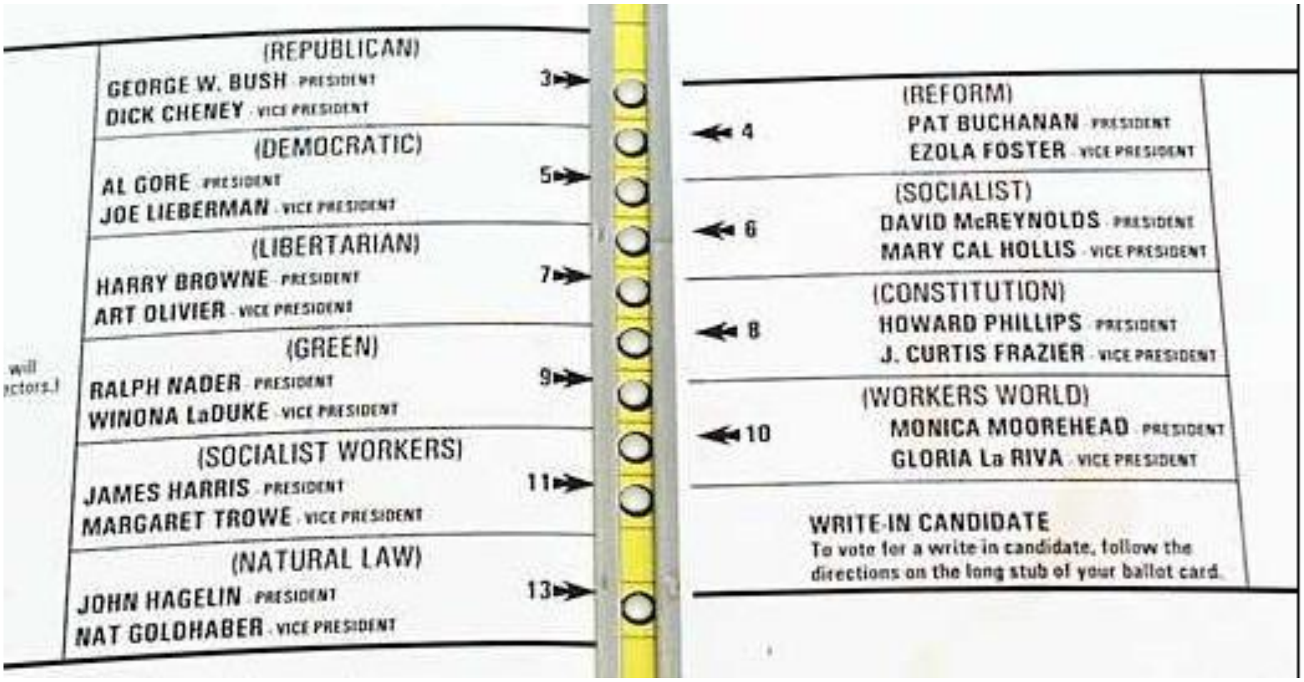


Figure 1

The aforementioned examples may not have occurred recently. However, technology will continue to evolve, and each new technology will come with its own challenges in design (Norman, 2013). Although the consequences of ambiguities are known, it is not known how ambiguities in products may be resolved. One possible way to understand ambiguous product designs is to examine whether they may be related with ambiguities in sentences, a topic that has been well studied in the linguistic literature. In understanding ambiguous products in conjunction with ambiguous sentences, we may also shed new light on the discussion about language relativity. Additionally, we examine how dual process theories and IQ may fit into this puzzle.

Ambiguous Designs

Many products in the world are confusing or unintuitive to use (Norman, 2013). We consider ambiguous product designs to be a subset of such confusing products. Ambiguous product designs are designs that where users have one goal when using the product, but the user may be inclined to have more than one interpretation of how the product should be used to realize the user's goal. However, due to the design of the product, there is only one way for the product to be used to achieve the goal that the user had. Ambiguous products can be correctly interpreted and used through interacting with the product. However, this need for a trial and error nevertheless makes the product more difficult and confusing to use than if there was no need for a trial and error. As mentioned previously, there are many products that are difficult and confusing to use, but that does not mean they fall into the category of ambiguous designs. The Norman door (figure 2) is not considered an ambiguous design for the purposes of this paper. This is because although the design is difficult to use (there is an intuition to push the door), there is a clear sign that says "pull", which removes the ambiguity.

Figure 2

An example of ambiguous design is the classic stove design in figure 2. It is unclear which of the dials control which of the burners in the image of the stove in figure 2. If the design was clear, it would obey Norman's (2013) rules of mapping, and they are as follows. 1.) Keep controls directly on what is being controlled. 2.) Controls are as close as possible to the object to be controlled. 3.) Controls are arranged in the same spatial configurations as the objects to be controlled. Using Norman's (2013) rules as a way of analysis, it is possible for dial 1 to control the red burner or blue burner. Similar ambiguities exist for the other dials. Hence, figure 3 shows how the stove is an ambiguous design.

Figure 3

There is a large literature about how designs can be made more effective and usable (Krug, 2014; Norman, 2013; Norman, 2004; Story, 1998; Nielsen, 1994). They describe principles that include creating consistent designs, enabling product discovery, and reducing potential error. Ambiguous designs especially violate principles of discoverability and error reduction. When considering the consequences of design flaws that were discussed in the introduction, these two principles are especially relevant. Hence, understanding ambiguous designs has valuable social and economic contributions.

However, there is not much known about how people interpret or interact with confusing or ambiguous designs that are difficult to use. Norman (2013) and Underhill (2009) provide evidence from careful observations of individuals. They suggest that people may give up and blame themselves when they encounter confusing design. However, they did not report any statistics to suggest how prevalent this is. Hence, this study also collects data to understand how people assign blame – do they blame themselves or do they blame the designer – when they encounter an ambiguous design.

Ambiguous Sentences

Product designs are not the only sources of ambiguities. Like ambiguous designs, sentences can oftentimes have more than one interpretation. There are many types of ambiguities that arise from language. They can arise because of multiple from lexical meanings. A classic example is “bank”, which has multiple definitions. When there is not enough context in a sentence or its environment, the multiple definitions of “bank” can allow for multiple interpretations of the sentence. Syntactical structures can also cause ambiguities. The following is an example of such an ambiguous sentence.

(1) I saw a man on a hill with a telescope.

Just as it is unclear which dial controls which burner in figure 1, it is unclear whether the phrase “with a telescope” is modifying “saw” or “the man”. However, there are sentences that are ambiguous in more subtle ways, such as:

(2) The bartender told the detective the suspect left the country yesterday.

In such a sentence, most readers prefer to associate “yesterday” with the action of the suspect leaving instead of the action of the bartender telling the detective something. Grammatically, the sentence can also be understood as the following. Yesterday, the bartender told the detective the suspect left the country. Most readers prefer the first reading because of a locality constraint, “yesterday” appears closer the act of leaving (Gibson and Pearlmutter, 1998).

There are other types of sentences that are like ambiguities but are not ambiguities. For example, vague sentences are similar to ambiguous sentences, because both types of sentences can have multiple meanings. However, vague sentences often have borderline cases. One instance is the word “short”. There is some cutoff for whether something is determined “short” or not. Even if the speaker and the listener of the sentence had access to the same information, there can still be disagreement over whether the object in question falls into the “short” category. In contrast, unique meanings in ambiguous sentences can be eliminated by context (Braun and Sider, 2007). If a speaker and a listener both have access to the same information, sentence 1 is no longer ambiguous. The ability for context to eliminate a possible interpretation of an ambiguous sentence is in parallel with ambiguous products. When a user can interact with an ambiguous product design, he/she will be able to eliminate one of the interpretations.

The parallels between ambiguous products and ambiguous sentences suggest that there may be a relationship between these two different domains. One cause to study these two together is that this relationship may be relevant for the ongoing interest in language relativity (discussed later). The other cause is that unlike the literature on product design, the linguistic literature offers a rich canon of theories regarding how people may interpret ambiguous sentences. This can help build on an understanding of what makes a design intuitive, something that Stangeland (2014) argues is still being developed.

In contrast to a sparse literature in resolving ambiguities in design, there is a rich literature on resolving ambiguities in sentences. Gibson and Pearlmutter (1998) present four possibilities for how constraints can influence the understanding of an ambiguous sentence. Those constraints are contextual, lexical, locality-based computational resource, and phrase-level contingent frequency. Spivey-Knowlton and Sedivy (1995) describe constraints that are specific to attachment ambiguities. Among the constraints, lexical context and environmental contexts are key to understanding ambiguous sentences (Linzen, Marantz, & Pylkkänen, 2013). In a sentence such as (1), there is no context when the sentence is by itself, which makes it an ambiguous sentence. If there is a relationship between ambiguity in product designs and ambiguity in sentences, these constraints from linguistics can be further explored to better understand the relationship between product designs and sentences. This can also have implications for how designs can be made less ambiguous and therefore increasing ease of use. As a parallel to ambiguous products, we also collect data to understand how people assign blame on ambiguous sentences – do they blame themselves or do they blame the speaker? (Blame is discussed later.)

Language Relativity

The Sapir-Whorf hypothesis or language relativity wrestles with how central language is to cognition. In other words, how are thought and language related? For example, is language a precursor for having thoughts? That possibility is strongly rejected for many reasons. One reason is that some thoughts, like tip-of-the-tongue phenomenon and imagery, occur without language (Casasanto, 2008). However, there are reasons to believe that language and thought are closely intertwined. One view is that language is a conduit for expressing beliefs about a wide variety of concepts (Carruthers, 2002). Some go further to claim that language augment humans' ability to understand object relationships, space, and number such that language enhances humans' capabilities in those domains (Gentner, 2016).

In reviewing several neurological studies, Thierry (2016) claims that language and thought are intrinsically linked together. There is no evidence that shows that there are language-specific regions in the human brain. Furthermore, "the spatial and temporal resolution of functional neuroimaging remains largely insufficient to establish any selectivity at the macrostructural level." (Thierry, 2016) Additionally, language has also been shown to influence visual and spatial processes, such as color perception. In their review of multiple empirical studies, Ünal and Papafragou (2016) argue that language influences the processing of colors. They write, "To summarize, studies addressing the interface between language and color processing reveal meaningful language-driven differences at the behavioral level. Furthermore, language influences appear at early stages of visual processing, as shown by ERP studies. Neuroimaging work also leaves open the possibility that even early stages of color processing might be susceptible to rapid linguistic feedback."

The effect of language is seen in other domains such as time and shapes and substances (Boroditsky, 2006). For example, people have a preference for how the order of photographs of a person at different stages of life should be arranged. Those that speak a language with a writing system that flows from left to right, like English, prefer to arrange the photos of the younger person on the left and the photos of the older person on the right. This is reversed for speakers of languages like Hebrew, where the language is written from right to left. However, Pormpuraawans, people part of an aboriginal group in Australia, do not use relative spatial terms like "left" and "right". Instead, they use absolute direction terms like "north" and "east". When they are asked to arrange the photographs of a person at different stages of life, they arrange the photos from east to west, regardless of whether the photos are arranged left to right or right to left (Boroditsky & Gaby, 2010).

The domains that language influence, such as visual and spatial processes, are no doubt related to the many mental processes that are activated when a person is interpreting a product design for use. However, we did not find any work in the literature that attempts to establish a link between those two domains. One example where language and design intersect is the location of the back button found on internet browsers and phones; to return to a previous page (as if they were going back in time), users would press a button that's usually located on the very left of the page with an arrow that points to the left, backwards in time. Taken together, this conglomeration of evidence suggests that ambiguities in product design and in language may be linked.

Cognitive Reflection

In examining an ambiguous product or sentence, a person may be recruiting System 1 and System 2 processes. According to Kahneman and Frederick (2002), System 1 is associated with rapid judgements and System 2 is a parallel and slower process that may endorse or refute the judgement of System 1. System 1 is generally considered to be instinctual, rapid, and associative, while Type 2 is generally considered to be deliberative and slow (Evans, 2003; Kahneman, 2011). The main difference between System 1 and System 2 (also referred to as Type 1 and Type 2) processing is that System 2 requires working memory and executive function (Evans & Stanovich, 2013). System 1, by contrast, tends to be automatic, and it does not require working memory. One description of System 1 processing is that it happens to a person. An example is facial recognition. When seeing an image, an individual automatically knows if that is the face of a person, and he/she can usually tell if the person in the image is a male or female, and all this happens with no conscious effort. By contrast, a System 2 task requires more effort, such as the task of multiplying 17 and 24 together (Kahneman, 2011). For most people, multiplying those two numbers would require working memory, a defining System 2 process, because the person would need to remember the product of 7 and 4, and then the product of 4 and 1, and so forth, so that the person can eventually add two numbers together to complete the original task.

When examining an ambiguous product or sentence, an individual may have an immediate intuition for what the correct interpretation is, which would be a System 1 reaction. Upon reflection, a System 2 process, that person may have a different idea, which may lead that person to want to change their answer or feel less confident about their choice. Participants that use this line of reasoning would use the cognitive abilities that the Cognitive Reflection Test (CRT) predicts.

The CRT is designed to simply and quickly measure "the ability or disposition to reflect on a question and resist reporting the first response that comes to mind." (Frederick, 2005) It has been shown to have correlations with many psychological traits, such as a preference to use System 2 to refute the judgements made by System 1 (Frederick, 2005).

Besides discerning an individual's preference for System 1 or System 2 thinking, the CRT has been shown to have other interesting correlates. Pennycook and Rand (2017) showed that good performance in the CRT is positively correlated with an ability to discern true news headlines from the fake ones. In another study, Pennycook, Cheyne, Barr, and Koehler (2015) discovered that the CRT can be used to predict the ability to detect pseudo-profound bullshit, "which consists of seemingly impressive assertions that are presented as true and meaningful but are actually vacuous." They argue that this is because the CRT has a demonstrated ability to select for participants based on their propensity for tasks that require System 2 processing (Campitelli & Gerrans 2013; Toplak, West, & Stanovich, 2011). In addition, the CRT has been linked to tasks associated with probability reasoning, under- and over- confidence, Wason selection task performance, and much more (Thomson & Opprenheimer, 2016).

In addition to the Frederick's (2005) CRT, we include an extended version of the CRT. It includes a question proposed by Toplak, West, and Stanovich (2014). They have found that their four-question version is comparable to Frederick's (2005) CRT in terms of predicting. When the two measures are combined, they find a substantial improvement in predicting cognitive ability, various thinking dispositions, and other rational thinking tasks. The other CRT that we also include was proposed by Thomson and Opprenheimer (2016). It appears to be less reliant on numeracy, unlike the original CRT. Additionally, we include other CRT items from working papers.

General Intelligence and IQ tests

Frederick (2005) demonstrates that the Wonderlic IQ test is positively correlated with the CRT. He explained that the two tests are likely correlated because being able to reach the correct answers in the CRT at least require mathematical ability. However, he noted that although the two tests are correlated, they measure different attributes.

When determining an interpretation to an ambiguous product or sentence, a person must first recognize that the product or sentence is ambiguous. This ability is correlated with a high CRT score. Once a person recognized the ambiguity, they can generate possible solutions and rationale for those possible solutions, which requires general cognitive ability. In the literature, general cognitive ability is referred to as "g", and according to the Catell-Horn model, it has two parts. One part is called fluid intelligence (gF) and it is associated with using mental processes to solve novel problems. For example, fluid intelligence is associated to drawing inferences, problem solving, extrapolation, identifying relations, and more. The other part is called crystallized intelligence (gC), and it is associated with a general breadth and depth of knowledge. Being able to name many different countries would be considered a having crystallized intelligence. (For more, see McGrew, 2005; Carroll, 1982; Cattell, 1963)

According to Hicks, Harrison, and Engle (2015), the Wonderlic test is able to discern between individuals with high and low fluid intelligence. In their study where participants completed the Wonderlic test along with other cognitive performances such as memory capacity, Hicks et al. (2015) demonstrate that the Wonderlic test can be used to predict fluid intelligence. For the exploratory purposes of this study, this falls in line with results from Bell, Matthews, Lassiter, and Leverett (2002) which shows that the Wonderlic test correlates with fluid intelligence. Overall, the literature on IQ is quite extensive and has been developed for many years with a general acceptance towards the Carttell-Horn model. Furthermore, the tests for IQ are generally accepted as having high statistical reliability, meaning that the scores generally agree with one another across time, even accounting for small variations (McGrew, 2005).

Locus of Control

Rotter's locus of control has been designed to measure how much internal control someone has over his/her life (Rotter, 1966). He describes locus of control as "a generalized attitude, belief, or expectancy regarding the nature of the causal relationship between one's own behavior and its consequences." One who has an internal locus of control believe that the events and outcomes in his/her life are largely a consequence of his/her action. In contrast, an individual with an external locus of control believe that outcomes are a result of causes outside of their control, such as others, luck, and the environments they encounter (Galvin, Randel, Collins & Johnson, 2018; Rotter, 1966).

Based on the research on locus of control, it appears be an appropriate measure to use when determining what psychological factors may be relevant in predicting how someone assigns blame when they are confused by an ambiguity in a product or a sentence. Although there are many different measurements of locus of control, the Rotter scale has been in long existence and it has been widely used. Thus, establishing a link with locus of control through the Rotter scale would allow us to look the many different domains that have used it. These domains include job satisfaction and life satisfaction (Wang, Bowling, & Eschleman, 2010). Therefore, we have decided to use it as a way to begin to explore blame in ambiguities in two different domains.

Hypothesis

Based on previous work showing the powerful influence of language on cognition, we believe that the ambiguities in product design and language are related; an ability to notice and differ ambiguity in one domain will predict an ability to do so in the other domain. This would be evidenced by a positive correlation between the response times for product ambiguities and response times for sentence ambiguities. Positive correlations between ambiguous products and ambiguous sentences would also exist in the participants' self-report of their confidence in their answers. Similarly, positive correlations should exist between the blame assignment for both ambiguous designs and ambiguous sentences.

We predict that CRT scores will positively correlate with response time. As participants weight the different likely options using System 2, they will spend more time to decide on an interpretation for each ambiguity. We do not believe that there will be a correlation between locus of control and response time as there is nothing in the literature to suggest there would be a correlation. We have two different hypotheses for how blame assignment will correlate with CRT and IQ. H1: Individuals with high CRT and high IQ will figure out that the design or sentence is flawed because of the designer or speaker, and they would blame the designer or speaker. H2: Individuals with high CRT and high IQ will think if they can't figure out an interpretation, they must have missed something, so they would blame themselves. Hypothesis 2 is much closer to Norman's (2013) characterization of blame assignment with regards to poor design. Finally, because locus of control is designed to measure belief about causality, those that lean towards an internal locus of control will be more likely to blame themselves, and those that lean towards an external locus of control will be more likely to blame others.

Method

We conducted a survey in Amazon Mechanical Turk, and we received responses from 355 respondents. Two participants were excluded, because they failed to pass basic English comprehension questions. These participants are not included in the rest of the paper. The participants consist of 213 men (60.3%) and 140 women (39.7%) with a mean age of 36.45. The standard deviation of age is 10.69.

Participants were presented with two sections; one section features a series of product ambiguities and the other section features a series of language ambiguities. Within each section of product ambiguities and language ambiguities, the order that the questions appeared in was randomized. Each section included control questions where ambiguity was minimized. Furthermore, for each ambiguity, participants were asked to choose one interpretation of the ambiguity in question. For example, participants are were to determine which dial control which burner in figure 3. For sentences, participants were also given two (or more) possible interpretations of the sentence (The appendix contains more examples). We chose our examples of designs based on real life designs that we have come across in the world, taking care to try to find a variety of examples where ambiguity arises due to poor mapping (such as the stove in figure 3), conflicting labels/cues (such as a fire exit sign with an arrow pointing one way, and a figure running from fire another way), and missing labels (such as a rocker switch without any context). Likewise, we included multiple examples of ambiguous sentences that arise due to lexical ambiguities (such as match) and syntactical ambiguities (such as attachment ambiguities mentioned in the ambiguous sentence section of this paper).

Then, they were asked to rate how confident they feel about their answer on a scale of 1 (corresponding to “no idea”) to 5 (corresponding to “certain”). At the end of each section, they are asked, “In the previous questions, how much do you blame yourself for any feeling of uncertainty?” Participants chose from a binary scale with 7 points, with the higher end corresponding to “I blame myself completely”, and the lower end corresponding to “I blame the designer[speaker] completely”.

After participants answer questions about ambiguity of products and sentences, they are complete a CRT, a shortened IQ test – Wonderlic Personnel Assessment and Raven’s Progressive Matrices (an often-used test for measuring IQ see Gignac (2015)), and Rotter’s locus of control questionnaire. Participants complete a demographics questionnaire at the end of the survey.

Results

We conducted correlation tests between the different variables. The results for each are summarized in the tables. In the tables that are shown, the number in each cell corresponds to the r-value of the correlation between the name of the row and the name of the column. Thus, in table 1, the starred r-value corresponds to the r-value between participants' response time of the control product questions and participants' response time of the control sentence questions. The highlighted cells correspond to correlations with p value less than .05. Additionally, each of the table that is presented below is part of a larger table that can be pieced together to show the correlations between all the variations. However, that table is too large to fit onto a single page.

Response Time

We do not find evidence that would suggest a correlation between the response rates to the product design control questions and the response rates to the sentence control questions. However, we find evidence that suggests that there is a positive correlation between response times (RT) of ambiguous designs and ambiguous sentences. This correlation is weaker than the correlation between the control designs RT and ambiguous sentences RT. Similarly, the correlation between control sentences RT and ambiguous sentences RT is also stronger than the correlation between ambiguous designs RT and ambiguous sentences RT.

	RESPONSE TIME		
	Amb Sent RT	Ctrl Design RT	Ctrl Sent RT
Amb Design RT	**.3**	**.47**	**.2**
Amb Sent RT		**.19**	**.37**
Ctrl Design RT			.08 ***

Table 1: Response times correlations

Key:

p:	<.05	<.01	<.001

Response Time Discussion

The lack of evidence for correlation between response times in product controls and sentence controls suggests the speed at which individuals process products and sentences are unrelated. On that basis, one may dismiss a relationship between product processing and sentence processing. However, there is evidence for a relationship between processing speed of ambiguous designs and sentences, based on their correlation ($r = .3$) with each other. Furthermore, this correlation ($r = .3$) is weaker than both the correlation between design response times ($r = .47$) and the correlation between sentence response times ($r = .37$). This can simply be because of an inherit relationship of mental processing within the same domain. But when we look at the correlations of ambiguous to control ($r = .19$ and $r = .2$) compared with correlations of RT of ambiguous items ($r = .3$), we see that these correlations between the controls and the ambiguous items are weaker than the correlation of the RT of the ambiguous items themselves. This suggests that there is a general relationship between processing of ambiguous items, but this relationship is weaker compared to the relationship of intra-domain processing as evidenced by $r=.47$ and $r=.37$.

One possibility is that there is some general executive process that resolves ambiguity in general. This would explain why we find evidence that individuals who are fast at processing ambiguous sentences are likely to be fast at processing ambiguous products. However, it appears that most of that correlation may be explained by speed in general, since fast processing of ambiguous items also correlates with fast processing of unambiguous items. Nevertheless, there is some correlation between processing speed of ambiguous items, which suggests that there may be some executive process that handles the processing of ambiguity. In considering the evidence, it appears that our data supports our hypothesis.

Confidence

We find a strong positive correlation (r = .6) between confidence in interpretations of ambiguous products and confidence in interpretations of ambiguous sentences. Other than evidence for negative correlation between confidence in control design RT and control sentences RT, there is no other evidence between response times and confidence scale. In general, we find that there is positive correlation across the board for the confidence scale. In other words, greater confidence in response to answers tend to correlate with greater confidence in other answers. The next strongest correlation is found between confidence within sentences.

		CONFIDENCE			
		Amb Design Conf	Amb Sent Conf	Ctrl Design Conf	Ctrl Sent Conf
Response Times	Amb Design RT	-.004	.07	-.02	.08
	Amb Sent RT	.04	-.003	-.01	.06
	Ctrl Design RT	.04	.01	**-.14**	-.02
	Ctrl Sent RT	.07	.01	.08	.004
Confidence	Amb Design Conf		**.6**	**.19**	**.27**
	Amb Sent Conf			**.27**	**.54**
	Ctrl Design Conf				**.39**

Table 2: Confidence correlations

Key:

p:	<.05	<.01	<.001

Although there is evidence for correlation in the confidence, a t-test suggests that the mean confidence for ambiguous designs and ambiguous sentences are different. The mean confidence for ambiguous designs is 2.17, and the mean confidence for ambiguous sentences is 2.75. The 95% confidence interval shows a difference between the means ranging from .488 to .687, the p-value is less than .001. In contrast, when we conduct a t-test for the difference in means between confidence in the control questions, the difference is smaller as shown by the

95% confidence interval (.018 to .20). The mean of confidence in control design questions is 3.51, whereas the mean of confidence in control sentence questions is 3.4.

	Mean	SD
Ambiguous Design Confidence	2.17	.67
Control Design Confidence	3.51	.55
Ambiguous Sentence Confidence	2.75	.68
Control Sentence Confidence	3.39	.68

Table 3: Mean & standard deviation of participants' confidence in their responses. Higher numbers correspond to self-reported confidence of participants' interpretations.

Confidence Discussion

It is probably not surprising to anyone that confidence in one domain leads to confidence in another. However, what is noteworthy is the difference in means of confidence in ambiguous designs and ambiguous sentences. This difference might be explained by the fact that humans use language more frequently and with more variety in their regular lives than they use product designs.

There is something different about the correlations in confidence versus the correlation in response times. Whereas the correlations in response times is better predicted within domains, the correlations in confidence is better predicted across the two different domains. In other words, as predicted by our hypothesis, being uncertain about ambiguity is a commonality that spans both product designs and sentences. Furthermore, there is something unique about ambiguity that causes the correlation between confidence measures of ambiguous items ($r = .6$) to be stronger than confidence measures where control items are used. That suggests there is some executive mechanism for processing ambiguities that is being shared by two different domains: designs and sentences.

CRT & IQ

We replicate the results from Frederick (2005) that demonstrated a correlation between the CRT and IQ scores. The correlation between Wonderlic scores and Raven's scores is in line with literature describing the statistical validity of IQ that was discussed in an earlier section. All of this establishes statistical validity of our survey.

The negative correlations we find suggest that individuals with a lower Wonderlic IQ score are taking more time to answer the questions. There is also evidence, albeit weaker, that individuals with lower CRT scores are answering the questions slower. However, this is not consistent with the Ravens IQ test, which do not appear to have correlation with RTs. In general, high IQ scores are correlated with decreased confidence in ambiguous items, but increased confidence in unambiguous items.

		CRT & IQ		
		CRT	Ravens IQ	Wonderlic IQ
Response Times	Amb Design RT	-.10	.07	**-.11**
	Amb Sent RT	**-.11**	-.004	**-.14**
	Ctrl Design RT	-.09	.08	**-.16**
	Ctrl Sent RT	**-.12**	.07	-.10
Confidence	Amb Design Conf	**-.18**	**-.22**	**-.36**
	Amb Sent Conf	-.09	**-.15**	**-.23**
	Ctrl Design Conf	**.13**	**.11**	**.25**
	Ctrl Sent Conf	-.02	.03	.03
CRT & IQ	CRT		**.28**	**.46**
	Ravens IQ			**.44**

Table 4: CRT & IQ correlations

Key:

p:	<.05	<.01	<.001

CRT & IQ Discussion

We predicted that a high CRT score would correlate with a longer response time for ambiguous items because individuals with a tendency to activate System 2 will be more likely to detect ambiguity. Our failure to find a positive correlation between CRT score and RT may be explained by two possibilities. One is that System 2 is often associated with, but not always associated with response times (Evans & Stanovich, 2013). This is made more likely by the other possibility, which is based on how we present participants with the questions. We provided participants with the other logical interpretations of the ambiguities; therefore, they did not have to consider generating their own, a process that is part of what the CRT measures. As such, individuals with greater reasoning abilities (as explained by general intelligence and measured with Wonderlic) have more confidence when they have reason to be confident (as in the case for answering control questions). Likewise, high IQ individuals have more reason to be less confident when they encounter an ambiguous question item; they are able to quickly correctly determine that both interpretations are “correct”. It is unsurprising that higher IQ scores are correlated with faster response times, because there was time limit in the administration of our IQ tests.

Our results also highlight an interesting difference between Raven’s IQ test and Wonderlic IQ test. Even though the two scores are correlated, only Wonderlic IQ scores correlate with response times. This may be attributed to the claim that Gignac (2015) makes in a literature review. He argues against the usage of the Raven’s IQ test as a measurement of general intelligence. Since the goal of this paper is not to address possible differences in IQ tests, this difference will not be further addressed.

Blame

We find no statistical significance that would suggest evidence for correlation between blame and locus of control. There is also no correlation between locus of control and any other variable we have presented.

Just as with confidence, blame tends to correlate with blame ($r = .49$). However, a t-test finds that the mean for blame in products (mean = 3.45, SD = 1.77) and the mean for blame in sentences (mean = 3.65, SD = 1.59) are likely the same ($p=.11 > .05$). In other words, people are attributing blame to both themselves and to the designer/speaker. The 95% confidence interval for the difference between those two means is -.45 to .05.

Higher confidence in answers to control (unambiguous) designs correlate to lower numbers on the blame scale (corresponding to a tendency to blame others). This pattern is not present in control sentences. However, higher confidence in answers to ambiguous items (design and sentences) correlate to higher numbers on the blame scale (corresponding to a tendency to blame the self). Overall, higher performance in CRT and IQ test is correlated with a preference to blame others (lower numbers on the blame scale).

Table 5: Blame. Lower numbers in blame represent participants attributing blame to others. Higher numbers represent participants attributing blame to themselves.

Key:

p:	<.05	<.01	<.001

		BLAME	
		Design Blame	Sent Blame
Response Times	Amb Design RT	.03	.06
	Amb Sent RT	.004	.02
	Ctrl Design RT	**.12**	**.11**
	Ctrl Sent RT	.10	.03
Confidence	Amb Design Conf	**.27**	**.13**
	Amb Sent Conf	**.14**	**.15**
	Ctrl Design Conf	**-.18**	**-.13**
	Ctrl Sent Conf	-.07	.0005
CRT & IQ	CRT	**-.19**	**-.15**
	Ravens IQ	**-.12**	**-.11**
	Wonderlic IQ	**-.32**	**-.30**
Blame	Design Blame		**.49**

Blame Discussion

One explanation for why we do not see evidence for a correlation between locus of control and blame assessment is that locus of control tends to measure granular factors in one's life. The meta-study that Wang, Bowling, and Eschleman (2010) conducts find that locus of control relates to work-related criteria like job satisfaction and burnout in addition to life satisfaction. Therefore, it would be reasonable to assume that locus of control measures at a larger granularity than more specific items like what we presented to participants in our study.

Our findings support our first hypothesis for how CRT and IQ may be related to blame assignment. The hypothesis was that individuals with high CRT and high IQ will figure out that the design or sentence is flawed because of the designer or speaker, and they would blame the designer or speaker. Seeing as how high IQ correlates to blaming others in ambiguous items but blaming the self in unambiguous item, we have evidence to support our hypothesis regarding blame. Additional evidence is present in how high IQ correlates with confidence negatively for ambiguous, but positively for unambiguous items. Norman (2013) suggest that people tend to blame themselves for failing to understand how a product design works. Our findings provide a more nuanced view. They suggest that people assign blame to both themselves and to designers, and that depends on the situation. Furthermore, how blame is assigned also varies depending on individuals' IQ.

General Discussion

An explanation for the correlation between the perception of ambiguities in products and in sentences is that there is an executive process in the mind that focuses on resolving all ambiguities. Across three measurements of perception of ambiguous (response time, confidence, and blame assessment), we find evidence that the two domains are related. Our results suggest that there is indeed some correlation between ambiguity in product designs and sentences. For example, there is evidence for correlation between processing speed of ambiguous items. This correlation cannot be fully explained by general processing speed (see discussion in an earlier section). In examining the other correlates this study explores, we believe that there is some other factor, such as a shared higher order executive process that is causing the correlation. This higher order executive process is likely related to IQ. There is evidence for positive correlation between ability to notice ambiguity and IQ (see earlier discussion regarding IQ and its correlation with confidence). Therefore, our evidence suggests that the executive processing is part of IQ tests (Wonderlic IQ in particular) measure.

Implications

It is hard to draw any strong conclusion on a study based on correlations as there can always be confounding variables. This evidence coincides with the existing literature about language relativity – language and cognition are related. As established in an earlier section, language relativity research shows that language is related to many facets of mental processing, including visual, spatial, and time processing (Gentner, 2016; Thierry, 2016; Ünal & Papafragou, 2016; Boroditsky & Gaby, 2010; Boroditsky, 2006).

Our data suggests that this higher order executive process is most likely an effect mediated by IQ rather cognitive reflection. The correlation between Wonderlic IQ and RT is stronger than the correlation between CRT and RT. We may have this finding because other possible interpretations of the ambiguous designs were already generated for participants to choose from. After all, a tendency for generating different possibilities is one of the measurements of the CRT. Therefore, our study should not be evidence against cognitive reflection playing a role in how individuals process ambiguities.

One implication for product designers is that words may give users a higher confidence in their interpretation of a designer. Attempts to use only illustrations or other visual cues to direct users may be misguided, especially if the illustration can be ambiguous. Thus, there is empirical evidence for how a designer might follow Norman's (2013) principles of design regarding discoverability. The viewpoint of including words would contradict those like Krug (2014). He suggests that less words make for better design. Rather, the viewpoint should be that less words can make for better design, but words are still helpful to increase users' confidence and improve their ability to correctly interpret the design. It should not be difficult to accept that increasing users' confidence in their interpretation of a design leads to good design. However, confidence in interpretation does not equal to good design. Based on our research, it is likely that people will interpret the Norman door correctly and confidently. But it is well known that the Norman door is an unintuitive design to use (Norman, 2013). And when users' confidence in ambiguous products is increased, they are more likely to blame themselves, which would decrease the likelihood of users providing feedback to designers, an essential component of improving design (Norman, 2013).

We do not discuss the specific responses to the ambiguous items (the subject of this is another paper). However, what appears to be happening is that certain ambiguous items correlate with higher IQ. In other words, although it may be the case that certain items have ambiguities that allow for multiple logical and valid interpretations, there is one interpretation that individuals with high IQ converge on. This suggests that although users may be converging on an interpretation of a design, this may be a result of users' IQ, not an intrinsic interpretability of the design. Furthermore, this study shows that even though people with higher IQs are converging on similar answers, the individuals have lower confidence in their interpretations of ambiguous items. Confidence about how to use a product design is undoubtedly related to visceral interactions with said product, and positive visceral interaction is linked to increased usability (Norman, 2004). Overall, if designers want to maximize the accessibility of their designs, they should strive to remove ambiguity. Otherwise, their design may only be usable to a subset of the population, which can have negative social and economic consequences depending on the environment of the design.

Limitations and Future Directions

Our data demonstrates there is evidence that suggests a correlation between processing of ambiguities in products and in sentences. As such, this study provides a new domain for which to understand the implications of the influence of language in cognition. This would be in line with previous work demonstrating the influence of language in spatial and visual domains (see earlier section), both of which are likely to be involved in the mental processing of product designs. We are not aware of any study attempting to link product design processing and language processing. As such, we hope that future research will continue to explore this potential link, which can have implications for established literature about language relativity.

One of the shortcomings of this study is that we do not follow the techniques often used by those that study language relativity. Studies regarding language relativity oftentimes use special populations, such as children and/or bilingual individuals. Participants also include speakers of different languages instead of only English speakers, which is what we did in this study.

The use of special equipment is another way that this new domain can be explored. For example, traditional linguistic techniques use eye movement tracking to observe how participants react when they are parsing an ambiguous sentence. Other techniques include the usage of brain scanning technology. Such techniques will be able to elucidate our understanding of the relation between ambiguous product designs and ambiguous sentences. There are undoubtedly many more methods to determine how ambiguity in designs and sentences (or language) may be related. Understanding this relationship will help provide another avenue to examine the language relativity problem. Ünal and Papafragou (2016) argue that “the question of whether language affects nonlinguistic cognition is too complex to be answered by a straight yes or no, and one needs to evaluate several fine-grained proposals to assess the specific conditions under which language interacts with cognition across different domains.” We propose that product designs may provide yet another domain to examine.

Conclusion

Designs have always been a part of human history, just like language. Here, we provide correlational studies to show how these two may be related. In doing so, we hope to further a discussion regarding language relativity. Furthermore, we hope that establishing a relationship between the two will help improve designs, thereby providing social and economic good. As designs become more and more complicated in a technology driven world, they will undoubtedly require higher order executive functions for humans to interpret and use them. Such designs should be easily accessible to everyone, regardless of their IQ.

References

Braun, D., & Sider, T. (2007). Vague, so untrue. *Noûs*, *41*(2), 133-156.

Bell, N. L., Matthews, T. D., Lassiter, K. S., & Leverett, J. P. (2002). Validity of the wonderlic personnel test as a measure of fluid or crystallized intelligence: Implications for career assessment. *North American Journal of Psychology*, *4*(1), 113-119. Retrieved from https://search.proquest.com/docview/197981450?accountid=15172

Boroditsky, L. (2006). Linguistic Relativity. In Encyclopedia of Cognitive Science, L. Nadel (Ed.). doi:10.1002/0470018860.s00567

Boroditsky, L., & Gaby, A. (2010). Remembrances of Times East: Absolute Spatial Representations of Time in an Australian Aboriginal Community. Psychological Science, 21(11), 1635–1639. https://doi.org/10.1177/0956797610386621

Carroll, J. B. (1982). The measurement of intelligence. In R. J. Sternberg (Ed.), Handbook of human intelligence Cambridge, England: Cambridge University Press.

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences*, *25*(6), 657-674. doi:10.1017/S0140525X02000122

Casasanto, D. (2008), Who's Afraid of the Big Bad Whorf? Crosslinguistic Differences in Temporal Language and Thought. Language Learning, 58: 63-79. doi:10.1111/j.1467-9922.2008.00462.x

Casey, S. M. (1998). *Set phasers on stun: And other true tales of design, technology, and human error*. Santa Barbara: Aegean.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1-22. doi: 10.1037/h0046743

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25-42. doi:10.1257/089533005775196732

Galvin, B. M., Randel, A. E., Collins, B. J., & Johnson, R. E. (2018). Changing the focus of locus (of control): A targeted review of the locus of control literature and agenda for future research. *Journal of Organizational Behavior*, *39*(7), 820-833.

Gentner, D. (2016). Language as cognitive tool kit: How language supports relational thought. *American Psychologist, 71*(8), 650-657. doi:10.1037/amp0000082

Gibson, E., & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences*, *2*(7), 262-268. doi:10.1016/s1364-6613(98)01187-5

Gignac, G. E. (2015). Raven's is not a pure measure of general intelligence: Implications for g factor theory and the brief measurement of g. *Intelligence*, *52*, 71-79.

Hicks, K. L., Harrison, T. L., & Engle, R. W. (2015). Wonderlic, working memory capacity, and fluid intelligence. *Intelligence, 50,* 186-195. doi:10.1016/j.intell.2015.03.005

Kahneman, D., & Frederick, S. (2002). Representativeness Revisited: Attribute Substitution in Intuitive Judgment. *Heuristics and Biases,* 49-81. doi:10.1017/cbo9780511808098.004

Krug, S. (2014). *Dont make me think, revisited: A common sense approach to Web usability*. Berkeley, CA: New Riders.

McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future.

Linzen, T., Marantz, A., & Pylkkänen, L. (2013). Syntactic context effects in visual word recognition: An MEG study. *Mental Lexicon*, *8*(2), 117–139. https://doi.org/10.1075/ml.8.2.01lin

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Celebrating Interdependence - CHI 94*. doi:10.1145/191666.191729

Norman, D. A. (2013). *The design of everyday things*. New York, NY: Basic Books.

Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. New York, NY: Basic Books.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied, 80*(1), 1-28. doi:10.1037/h0092976

Spivey-Knowlton, M., & Sedivy, J. C. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition, 55*(3), 227-267. doi:10.1016/0010-0277(94)00647-4

Stangeland, A.M. (2014). Intuitive use in design guidelines Can intuitive use be applied to a product ?

Story, M. F. (1998). Maximizing Usability: The Principles of Universal Design. *Assistive Technology, 10*(1), 4-12. doi:10.1080/10400435.1998.10131955

Thierry, G. (2016). Neurolinguistic Relativity: How Language Flexes Human Perception and Cognition. *Language Learning, 66*(3), 690-713. doi:10.1111/lang.12186

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision making, 11*(1), 99.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, *20*(2), 147-168.

Ünal, E., & Papafragou, A. (2016). Interactions Between Language and Mental Representations. *Language Learning*, *66*(3), 554-580. doi:10.1111/lang.12188

Underhill, P. (2009). *Why we buy: The science of shopping*. New York, NY: Simon & Schuster.

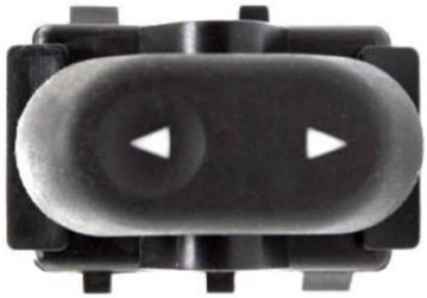
Wand, J. (2001). The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida. *American Political Science Review, 95*(4), 793-810.

Wang, Q., Bowling, N. A., & Eschleman, K. J. (2010). A meta-analytic examination of work and general locus of control. *Journal of Applied Psychology*, *95*(4), 761.

Appendix

Ambiguous Designs Questions

Below is a control for power windows in a car. Which rolls the window **UP**?



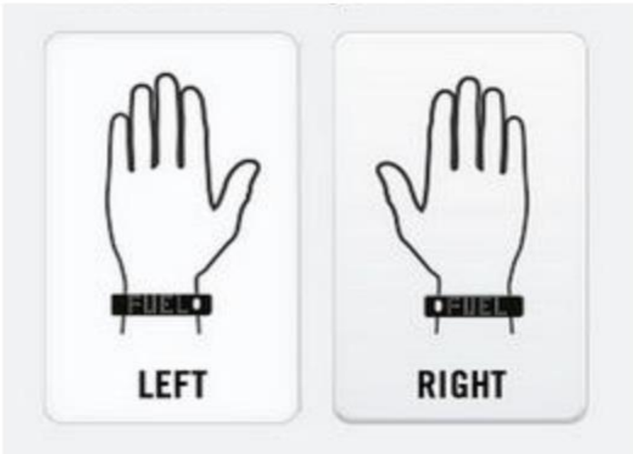
Pressing ◁	Pressing ▷

Which button do you press to go up?

🔴	🔵

Nike's Fuel Band is a wearable activity tracker (like a Fitbit). To set up the device, you must select the wrist you'll wear it on.

For which wrist is the Fuel Band currently selected?



Left	Right

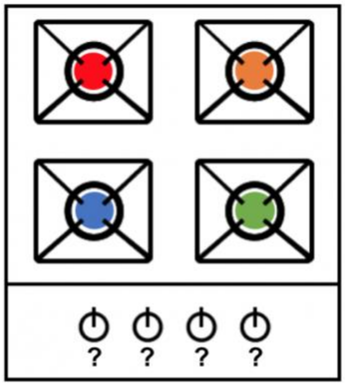
Which way should you run if you smell smoke?

Left	Right

Select all of the burners that are on, according to the display.




Going from leftmost dial to rightmost dial, indicate which burners they should control.



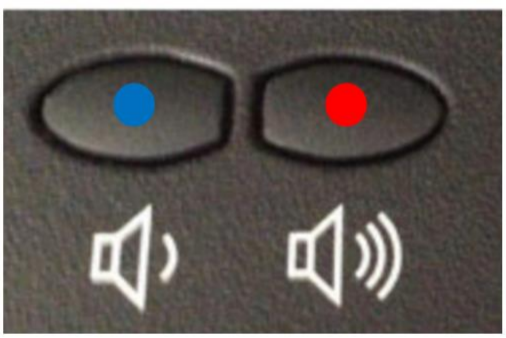
Other

Unambiguous Design Questions used as Control

The picture shows the buttons next to a car seat.

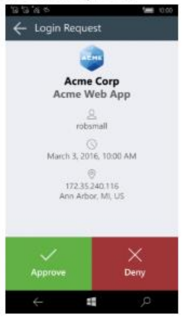
If you want to adjust the headrest, which button would you press?

Supposed you saw the following buttons next to a speaker.



Which colored button would you press to decrease the volume?

Which button do you press to approve the request?



Left

Right

Ambiguous Sentence Questions

Suppose you received notice that a meeting that was originally scheduled for **next Wednesday** has been **moved forward two days**.

When would you expect the meeting to occur?

Monday

Friday

On **Monday, April 15,** Jim told Nathan that he'd be available to meet **next Saturday**.

APRIL 2019

SUN	MON	TUE	WED	THU	FRI	SAT
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

Download & Print Free Calendars From www.wiki-calendar.com

What day should Nathan keep clear on his calendar?

Saturday, April **20**

Saturday, April **27**

I saw a man on a hill with a telescope.

There's a man on a hill, and I'm watching him with my telescope.

There's a man on a hill, who I'm seeing, and he has a telescope.

Look at the dog with one eye.

Look at the dog using only one of your eyes.

Look at the dog that only has one eye.

We invited the performers, Danzel and Desha.

We invited the performers, and we invited Danzel, and we invited Desha.

We invited the performers named "Danzel" and "Desha."

The bartender told the detective that the suspect left the country yesterday.

Yesterday, the bartender told the detective that the suspect left the country.

The bartender told the detective that yesterday, the suspect left the country.

Some girl read every book.

There's a specific girl who has read all the books.

For each book, there's some girl who has read it.

Let's stop controlling people.

Let's stop people who control others.

Let's stop controlling other people.

She is looking for a match.

She wants to start a fire.

She is searching for the other half of a pair.

Visiting relatives can be exhausting.

It is exhausting when relatives visit you.

It is exhausting when you go visit relatives.

Unambiguous Sentences used as control

Look at the man that has binoculars.

Look at the man by using binoculars.

Look at the man who has binoculars.

On **Monday, April 8,** Al told Jamie that he'd be available to meet **Saturday**.

APRIL 2019

SUN	MON	TUE	WED	THU	FRI	SAT
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

What day should Jamie keep clear on his calendar?

Saturday, April **13**

Saturday, April **20**