DIVISION OF COMPUTER SCIENCE

Seven Lesser Known Myths of Formal Methods: Uncovering the Psychology of Formal Specification

R J Vinter, M J Loomes and D E Kornbrot

Technical Report No. 250

April 1996

Seven Lesser Known Myths of Formal Methods: Uncovering the Psychology of Formal Specification

Rick Vinter, Martin Loomes and Diana Kornbrot
School of Information Sciences
(in collaboration with the School of Health and Human Sciences)
University of Hertfordshire, Hatfield, Hertfordshire, UK

Abstract

Psychological research has shown that people are prone to systematic errors when reasoning about logical statements in natural language. The Human Cognition and Formal Methods research project aims to test whether people are equally susceptible to error when reasoning about the same types of logical statement in formal notations. A series of specially designed experiments plan to investigate specific properties of formal notations that could affect the ease with which people are able to understand and reason about formal specifications. The first experiment concentrated on five cognitive activities which are central to the process of developing a formal specification: reading, writing, understanding, translating and reasoning. It also examined the ways in which a designer's writing style can affect his or her audience's understanding of a specification. The results of the experiment suggested that some of the software engineering community's widely held beliefs about formal methods might, in fact, be misconceptions. This paper uncovers seven such "myths" based on the experiment's findings and discusses their possible implications for the future practice of software specification.

Introduction

> "Rather than seeing failure and errors as things that exist, but can be avoided with the right methodology, we can view them as things that the designer brings about, and ask what behaviour causes this. If we better understood why designers make mistakes we might be able to suggest ways they can adjust their behaviour to minimise errors, or contain their impact on the process as a whole."
>
> Loomes, Ridley and Kornbrot [Loo94].

Cognitive psychology is that branch of psychology concerned with explaining the ways in which humans store, manipulate and use information. Over the past three decades, cognitive psychology has made considerable progress towards understanding some of the inherent problems that people appear to have when reasoning about certain kinds of logical statement in natural language. Previous studies have found that people are prone to systematic errors when reasoning about disjunctive [Gri84], conjunctive [Lak71], negative [Joh72] and conditional [Bra91] statements. The Human Cognition and Formal Methods research project aims to test whether some of these findings carry over into the domain of formal specification by conducting a series of specially designed experiments. There are legitimate reasons for suspecting that the same patterns of errors do carry over because most formal notations contain grammatical constructs whose formal semantics correspond to those of their natural language counterparts ($\vee$, $\wedge$, $\neg$, $\Rightarrow$). If formal specifications can be shown to incite the same forms of reasoning errors, this should constitute a genuine reason for concern because developers who are unable to interpret or reason clearly about system specifications are more likely to make the types of erroneous development decisions which, in the past, have led to the production of defective systems. Many of the processes associated with formal specification are complex and loosely defined which means that their potential for human error is high. Since formal methods are commonly used in the development of safety-critical systems, attempts to minimise this potential for error should be welcomed.

The emergence of formal methods has also seen the emergence of numerous misconceptions regarding their intended purpose, compositions, prerequisites and commercial viability. In two highly influential papers written by Hall [Hal90] and Bowen and Hinchey [Bow94], fourteen common "myths" or misconceptions about formal methods were uncovered. This paper proposes a further seven, lesser known myths relating specifically to the psychology of formal specification. Some of the myths described are favourable to formal methods; some are unfavourable. However, all are supported by empirical evidence generated by an initial investigation conducted at the University of Hertfordshire during the latter half of 1995. Its participants comprised six university staff and six students, all of whom possessed computer science backgrounds and had some prior knowledge of the Z notation [Spi92]. The experiment comprised four tasks divided into a total of ten parts.

Seven Lesser Known Myths

Myth 1. *People always reason logically about formal logic based specifications.*

> "There is a certain degree of 'trade-off' between the expressiveness of a language and the levels of abstraction that it supports. Making a language more expressive does indeed facilitate briefer and more elegant specifications, but it can make reasoning more difficult."
>
> Bowen and Hinchey [Bow94].

Perhaps one of the main advantages of formal logic over natural language is that it abstracts away extraneous information and allows reasoners to concentrate purely on the underlying form of arguments. It therefore seems intuitive that, in general, it would be easier to reason about a system's description in a formal notation than the equivalent description expressed in a natural language. However, results from two separate tasks in the first experiment suggested that people's judgements are more likely to be logically valid when reasoning about specifications expressed in natural language than in a formal notation. Furthermore, they suggested that people do not always adhere to logical rules when reasoning about formal specifications, even when such rules are well known to them.

Wason's Abstract Selection Task

$\boxed{A}$ $\boxed{4}$ $\boxed{S}$ $\boxed{7}$

(A) (B) (C) (D)

If there is an A on one side of the card then there is a 4 on the other.

Which cards would you need to turn over in order to determine whether the rule is true or false?

Task 1: The Formalised Selection Task

InOut
$in? : Letter$
$out! : \mathbb{N}$

$(in? = A) \Rightarrow (out! = 4)$

(A) $in? = A$ (B) $out! = 4$
(C) $in? = S$ (D) $out! = 7$

Which inputs and outputs would enable you to test whether 'InOut' is working correctly?

The first experimental task presented participants with a variation on the Wason abstract selection task [Eva93, p. 99-135; Was66] set within the context of a Z formal specification. Wason's selection task is one of hypothesis testing and deductive reasoning based on conditional logic. The choices of responses shown to participants corresponded to the p, q, $\neg p$ and $\neg q$ cases for a conditional rule of the form *if p then q*. In Wason's standard version of the task, the conditional rule *if p then q* is implicit, whereas in the formalised version, the conditional was shown in the form of an explicit logical implication statement. Participants need to employ both the affirming *modus ponens* (MP) and denying *modus tollens* (MT) inference rules in order to deduce the correct combination of responses: namely, the cases corresponding to p and $\neg q$. Intuitively, it seems reasonable that those people with a background in formal logic would be more likely to recognise the type of mental inference required in order to deduce the correct response because their recognition of the explicit logical implication operator would ensure that they endorse only what follows logically. So, prior to the experiment, a higher rate of correct selections was predicted than the 4% observed during Wason's early trials [Was72, p. 182]. However, the actual success rate of 0% for the formalised task came as somewhat of a surprise. The observed combinations of responses suggested that all participants had successfully applied the MP form of inference but few, if any, had evaluated the MT form as being relevant. One explanation for the high rate of erroneous responses in both versions of the task is that participants had succumbed to a form of "matching bias," whereby they focused on those terms explicitly mentioned in the conditional rule [Eva83]. These participants' selections were therefore based mainly on probablistic guesswork rather than logical deduction. Clear correlations between the results from Wason's abstract version and the formalised version of the selection task suggested that Wason's findings do indeed carry over into the domain of formal specification and that, contrary to intuition, people do not necessarily find it easier to reason about conditionals in formal logic.

Task 4: A Summary of the Correct Syllogistic Inferences

Modus ponens:
$(shape = circle) \Rightarrow (colour = blue)$
$shape = circle$
Therefore, $colour = blue$

Modus tollens:
$(shape = circle) \Rightarrow (colour = blue)$
$colour = red$
Therefore, $shape \neq circle$

Modus tollens (negative antecedent):
$\neg(shape = circle) \Rightarrow (colour = blue)$
$colour \neq blue$
Therefore, $shape = circle$

Denial of antecedent avoided:
$(shape = triangle) \Rightarrow (colour = red)$
$shape = square$
Therefore, nothing follows

Affirmation of consequent avoided:
$(shape = square) \Rightarrow (colour = green)$
$colour = green$
Therefore, nothing follows

An investigation conducted by Evans [Eva77] aimed to determine whether the linguistic form in which arguments are presented and the presence or absence of negative components affect the rates at which people are able to draw valid inferences from given premisses. His results suggest that the rates at which participants drew successful inferences and succumbed to classical fallacies when reasoning about conditional syllogisms in natural language could be lowered or raised significantly by manipulating these independent variables. For each part in the present experiment's fourth task, participants were presented with two syllogistic premisses in the form of Z predicate expressions: one conditional and one equivalence. In each case, participants were required to draw a different kind of inference in order to arrive at a logical conclusion (as shown above). The purpose of the task was to determine whether presenting conditional syllogisms in formal logic would affect the rates at which people drew correct inferences or succumbed to reasoning fallacies. The observed results suggested that, although every participant drew the simple MP inference, only one third made the simple and more complicated forms of MT inference. This latter result indicated that participants were not at all distracted by the presence or absence of the negative operator. The results also indicated that most participants had avoided committing the two classical reasoning fallacies: denying the antecedent and affirming the consequent. The results from the fourth task were then compared with those obtained during Evans' natural language based study. Firstly, this between studies comparison suggested that people are much less prone to committing the two reasoning fallacies when reasoning about formal logic rather than natural language. Secondly, it suggested that, although people are equally adept at drawing MP inferences in both linguistic forms, people find it much more difficult to draw valid MT inferences in formal logic than in natural language. The fact that such a large proportion of participants failed to make the MT inference might begin to explain the same participants' poor performance on the formalised Wason selection task, where it was necessary to draw both MP and MT inferences in order to see the relevance of the correct responses.

In the past, psychology has pointed to convincing evidence which suggests that people frequently stray from what logically follows when reasoning about arguments expressed in natural language [Byr89; Eva93]. This begs the question which was a major concern of the first experiment: does people's reasoning conform more closely to the rules of formal logic when they are reasoning about formal logic itself? Intuitively, one would think so, but results from the first experiment suggested otherwise. Firstly, every participant's response to the formalised Wason selection task was illogical. That is, their selected combinations of inputs and outputs would not have enabled them to deduce for absolute certainty whether the rule was true or false. Although every participant appeared to correctly evaluate the p case as being relevant, none appeared to see the relevance of the $\neg q$ case. One might postulate that, if participants had known how to perform the MT form of reasoning needed to identify the $\neg q$ case as being relevant, then they would have deduced the correct response. However, results from the formal syllogistic reasoning exercises indicated that at least one third did know how to perform both the simple and complicated forms of MT reasoning. So, although many participants may have possessed MT in their mental repertoires of inference rules, none actually identified it as being applicable to the formalised selection task. The question of why the same number of participants did not derive the correct responses for both tasks can perhaps be answered by the fact that people's deductive performance rarely equals their deductive competence and the possibility that performance can be impaired or facilitated merely by changing the way in which a problem is presented. Overall, the results obtained from the first and fourth task constitute evidence that people do not necessarily find it easier to reason logically about explicit conditionals in formal logic than implicit conditionals in natural language.

Myth 2. *Most readers find precise specifications the easiest to understand.*

> "A succinct formulation of a certain property may seem adequate to one; while another will prefer a more verbose exposition of its consequences. To communicate clearly with the majority of readers, you should, in general, prefer clarity to brevity."
>
> Gravell [Gra91].

Gravell's claim, based on an informal "straw poll" of software engineers' opinions, suggests that audiences are more likely to understand clear (i.e. precise) specifications rather than brief (i.e. concise) specifications. Intuitively, it seems fair to assume that a clearly written and more detailed specification, as opposed to a brief and abstract one, would be more clearly understood by a majority of its audience. The main aims of the third experimental task were to discover participants' writing style preferences and, at the same time, to test Gravell's claim empirically.

Task 3: The Four Specification Styles

$$SWITCH ::= on \mid off$$

$$\begin{array}{|l}
\multicolumn{1}{l}{\textit{Toggle}} \\ \hline
s, s' : SWITCH \\ \hline
s' \neq s \\ \hline
\end{array}$$

Concise

$$\begin{array}{|l}
\multicolumn{1}{l}{\textit{Toggle}} \\ \hline
s, s' : SWITCH \\ \hline
(s = off \land s' = on) \lor \\
(s = on \land s' = off) \\ \hline
\end{array}$$

Verbose

$$\begin{array}{|l}
\multicolumn{1}{l}{\textit{Toggle}} \\ \hline
s, s' : SWITCH \\ \hline
s = on \Rightarrow s' = off \\
s = off \Rightarrow s' = on \\ \hline
\end{array}$$

Precise

$$\begin{array}{|l}
\multicolumn{1}{l}{\textit{Toggle}} \\ \hline
s, s' : SWITCH \\ \hline
(s = on \lor s = off) \Rightarrow \\
(s' = on \lor s' = off) \\ \hline
\end{array}$$

Imprecise

Participants were shown an English description of a software operation and four different Z implementations: one concise, one verbose, one precise and one imprecise. They were asked to select which style "best describes" the operation's behaviour and to justify their selections appropriately. The results from this task suggested that participants held equal preferences for each of the concise, verbose and precise styles, but held a universal dislike of the imprecise version. So, despite his suggestion that precise specifications are highly desirable, the observed results actually ran contrary to Gravell's claim. They also exhibited strong correlations between participants' ages, experience and preferred specification styles. Whilst the youngest and least experienced tended to choose the concise style, the oldest and most experienced appeared to prefer the precise style. Overall, these results imply that precision is not universally desirable and that, in order to communicate effectively with a majority of readers, designers must carefully consider the type of audience for which they are writing.

Considerate designers writing for novice readers might aim to specify the maximum amount of detail clearly so as to leave nothing to chance, using only the simplest notational constructs.[1] This might enable all of a document's potential audience to comprehend, without relying upon readers' knowledge of the notation's more complex features, but at the expense of expert readers finding the document more laborious to read. In contrast, considerate designers writing for an expert audience might aim to specify the minimum detail necessary by freely using the full range of a notation's constructs, leaving readers to infer for themselves the other implicit properties of system functionality. In this case, designers rely entirely upon their audience's expert knowledge of the notation. Here, there is always a danger

[1] The terms "expert" and "novice" are used here to distinguish between those readers who do and do not possess full knowledge of a notation's grammatical rules and constructs, respectively.

that novice readers will not be able to comprehend certain parts of the specification and will accept the first plausible meaning that appeals to their intuitions, as exemplified by participants' responses for the Z to English translation task. This may be dangerous because readers might use their inaccurate interpretations as a false basis from which to make incorrect judgements.

The extent to which a particular writing style coincides with a reader's natural form of interpretation might depend upon numerous independent variables which add further implicit meaning to what is said explicitly. These include: what is being specified and its surrounding context, the reader's prior knowledge and their language expertise. In theory, it might be argued that the ideal level of abstraction that one could use in a specification would take into account both its audience's prior knowledge and their language expertise. However, in practice, specifications are normally aimed at different readers with differing backgrounds. It is obviously impractical for designers to write several versions, each one aimed at a particular group with a certain level of expertise. This might explain why precision is rarely compromised in reality and the maximum amount of detail is always stated explicitly so as to leave nothing to chance. Whether this principle should be applied in all cases is debatable.

Clearly, in the case of the concise *Toggle* specification presented during the third task, most people are immediately able to deduce that, if the switch's setting is *on* before the operation is executed then it must be *off* afterwards. They are able to infer this immediately and without recourse to the *SWITCH* type declaration because their prior knowledge of electronic devices tells them that switches normally have two opposing states, *on* and *off*. However, it might be argued that the degree of concision shown in the first specification would not be suitable for use in certain applications, nor indeed suitable for certain kinds of audience.

$$COLOUR ::= red \mid green$$

$$\begin{array}{|l}
\multicolumn{1}{l}{ToggleColour} \\ \hline
c, c' : COLOUR \\ \hline
c = red \Rightarrow c' = green \\
c = green \Rightarrow c' = red \\ \hline
\end{array}$$

Precise

$$\begin{array}{|l}
\multicolumn{1}{l}{ToggleColour} \\ \hline
c, c' : COLOUR \\ \hline
c' \neq c \\ \hline
\end{array}$$

Concise

The two specifications of operation *ToggleColour* (shown above) illustrate one situation where precision might be used in preference to concision. Although the first exhibits exactly the same level of concision as that contained in the original *Toggle* operation, readers cannot rely upon natural intuition alone in order to deduce this operation's post-condition; this time, the *COLOUR* data type declaration must be referenced. For this kind of application, then, the precise style appears more suitable. Of course, in this instance, the savings in reading time and effort are negligible, but for realistic large-scale applications, where literally hundreds of mental references might be avoided in a similar fashion, a document's overall readability can be enhanced considerably. From this, it can be concluded that it is indeed possible for concision to be used with great effect in specifications, however, its effectiveness will depend largely upon readers' prior knowledge of what is being specified; what one person may consider trivial and take for granted, another may require further explicit elucidation.

Myth 3. *A formal specification is a precise and unambiguous description.*

> "When specifications are used as a communication medium among programmers during a system design and implementation, it is essential that programmers reading a specification all agree on what the specification means. This is more likely when the specification is formal, for two reasons. First, there is only one way to interpret a formal specification, because of the well defined and unambiguous semantics of the specification language. Second, the formality of the language encourages greater rigor in the definitions."
>
> Liskov and Berzins [Lis79, p. 278].

In view of the fact that formal languages are based upon grammatical rules and definitions, one might be forgiven for thinking that people always interpret formal specifications according to these alone. After all, this is how a machine would reach an interpretation. Indeed, it is perhaps this belief which led Liskov and Berzins, like much of the software engineering community, to assert that "there is only one way to interpret a formal specification". But people are not machines and the initial study produced overwhelming evidence of this. Its results suggest that people are liable to interpret formal specifications according to their own methods which can give rise to interpretations that do not coincide with the logical meanings prescribed by the notation's underlying formal semantics. In other words, formal specifications are not generally understood solely according to what is explicitly stated within them and people's prior knowledge can significantly bias their interpretations. Participants' responses for the Z to English translation task suggested that, rather than attempt to deduce the precise meaning of its most complex expression, they abandoned their knowledge of the Z notation and reverted to their own heuristic methods in order to form a plausible, but incorrect, interpretation. Specifically, it is thought that all participants obtained the gist of the expression's meaning by scrutinising only its key linguistic components (i.e. its variable identifiers) and relating these to their own misleading preconceptions of the type of application being specified.

It is a serious misconception for anyone to think that a specification must be precise and unambiguous simply because it is expressed in a formal notation. The problem is that there exist certain types of ambiguity that formal grammars cannot prevent from arising; it is the task of software designers to minimise the number of ways in which their specifications can be interpreted. The imprecise specification of operation *Toggle* presented during the third task illustrates one such form of ambiguity. Although it is written in valid Z notation, it does not define the relations between its input and output variables s and s' in sufficient detail for readers to predict the operation's post-condition with absolute certainty, given its pre-condition. Liskov and Berzins emphasise the importance of development staff agreeing on a single interpretation of a system's specification. However, whether a formal specification would be any more likely to result in this happening than other forms of communication remains to be proven scientifically.

Myth 4. *Formal expressions map clearly, uniquely and intuitively into equivalent natural language statements, and vice versa.*

> "Providing a bilingual person with information in one of his languages and testing him for it in the other enables us to study how the mind handles different kinds of information. It also enables us to separate skills in handling information from the content or information itself."
>
> Kolers [Kol73].

Given a text written in a foreign or technical language, it is generally believed that readers will implicitly attempt to translate each part into an appropriate form in their native languages before attempting to reason about its contents. In this light, it seems important that formal expressions can be converted to and from equivalent natural language forms clearly and intuitively so as to cause minimum distraction when people are reasoning about formal specifications. However, results from the first experiment suggested that, in practice, the mapping is often far from being clear and intuitive and that, more alarmingly, significant properties of formal expressions can be lost during this implicit conversion process.

Task 2a: The Modified Library System Specification

Library

$stock : Copy \mathrel{+\!\!\!\!\!\rightarrow} Book$

$issued : Copy \mathrel{+\!\!\!\!\!\rightarrow} Reader$

$shelved : \mathbb{F}\, Copy$

$readers : \mathbb{F}\, Reader$

$shelved \cup \operatorname{dom} issued = \operatorname{dom} stock$

$shelved \cap \operatorname{dom} issued = \varnothing$

$\operatorname{ran} issued \subseteq readers$

$\neg \exists\, r : readers \bullet \neg(\#(issued \rhd \{r\}) > maxloans)$

Original fourth predicate: $\forall\, r : readers \bullet \#(issued \rhd \{r\}) \leq maxloans$
The number of books that any reader borrows must be less than or equal to the maximum number of loans allowed.

Revised fourth predicate: $\neg \exists\, r : readers \bullet \neg(\#(issued \rhd \{r\}) > maxloans)$
The number of books that any reader borrows must be more than the maximum number of loans allowed.

The first part of the second experimental task presented participants with a formal specification originally presented by Potter *et al.* [Pot91, p. 124] for describing the abstract state of a computerised library system. However, for purposes of this experiment, the fourth predicate was modified to oppose people's general conceptions of library systems (described above). Participants were asked to translate the specification's predicate part into an appropriate form in natural English. Although most were able to offer translations which preserved the original meanings of the first three predicates, the fact that no participants gave consistent translations of the fourth expression suggested that significant properties of specifications can indeed be lost during translation to natural English. The form of their translations suggested that, instead of deducing the predicate's meaning from its grammatical constructs, participants relied solely upon their misleading prior experience of library systems in order to arrive at their interpretations. That is, all participants appeared to use probablistic inferences about possible relations between the fourth predicate's key linguistical components (i.e. its variable identifiers) and the surrounding context in order to arrive at plausible, but incorrect, translations. Above all, the results from this exercise stress that formal specifications sometimes do not have intuitive corresponding natural language translations and that interpretational bias can be caused by people's prior knowledge of the same or similar kinds of application to the one under specification.

Task 2b: Two Possible Z Implementations

Operation 'ComputeValue' outputs the sum of its two inputs squared.

$$\begin{array}{|l}
\multicolumn{1}{l}{\_\_\ \mathit{ComputeValue}\ \_\_\_\_\_\_} \\
\mathit{in}1?, \mathit{in}2? : \mathbb{Z} \\
\mathit{out}! : \mathbb{Z} \\
\hline
\mathit{out}! = (\mathit{in}1? \times \mathit{in}1?) + \\
\qquad (\mathit{in}2? \times \mathit{in}2?) \\
\hline
\end{array}$$

Solution A

$$\begin{array}{|l}
\multicolumn{1}{l}{\_\_\ \mathit{ComputeValue}\ \_\_\_\_\_\_} \\
\mathit{in}1?, \mathit{in}2? : \mathbb{Z} \\
\mathit{out}! : \mathbb{Z} \\
\hline
\mathit{out}! = (\mathit{in}1? + \mathit{in}2?) \times \\
\qquad (\mathit{in}1? + \mathit{in}2?) \\
\hline
\end{array}$$

Solution B

The second task required participants to translate a natural English requirements description into the Z notation. The description shown was open to two possible interpretations and it was possible for participants to have offered two corresponding Z implementations which described very different operations, both of which were nonetheless consistent with the ambiguous requirements description (shown above). Initially, it was predicted that their knowledge of elementary mathematical principles would lead most participants to offer implementations resembling the form of solution A because the rules of arithmetic state that multiplication precedes addition wherever there is an absence of parentheses. However, the responses to this task indicated that opinions were equally divided over the two possible forms of solution. Since it is important that a specification conveys a clear and unique message to every member of a development team, this exercise demonstrated the inadequacy of natural language for expressing requirements specifications insofar as it is prone to imprecise and ambiguous interpretations. It also emphasises the need for designers to exercise caution when contemplating which aspects of their audience's prior knowledge are taken for granted. Discussion of the results from each of the second task's two parts has thus far shown that formal expressions do not always translate clearly and intuitively into natural English, and *vice versa.* It has also revealed that significant properties of specifications can actually be lost during the translation process. Furthermore, the fact that no two participants gave exactly the same solutions for either of these two exercises suggests that there rarely exists a unique, one-to-one correspondence between formal and natural language statements.

Myth 5. *Writing a formal specification is a systematic process.*

> "There is a fundamental logical objection to verification, an objection on its own grounds of formalistic rigor. Since the requirement for a program is informal and the program is formal, there must be a transition, and the transition itself must necessarily be informal."
>
> DeMillo, Lipton and Perlis [DeM79].

In the past, the term "formal methods" has been associated with both the processes involved in writing specifications and the languages used to express them. Whilst the latter are nearly always formally defined, the former are not; the process of writing a formal specification rarely comprises any explicitly predefined sequences of actions whatsoever. One might claim that the advent of automated machine checking and animation has formalised the verification process, yet designers still play a major role in deciding which parts of a specification are to be proven and how the proofs are actually carried out. So, in a similar manner to that described by DeMillo *et al.* for the transition from informal requirements to formal program code, the transition from a set of informal requirements to a formal specification must necessarily be informal. From this perspective, then, the term "formal methods" appears to be somewhat of a misnomer. From it appears to have sprung a common misconception that formal specifications are produced via some systematic means, whereby designers exert little control over how specifications are developed and the style in which they are eventually written. Of course, if this were true then, given the same set of requirements, any number of designers could follow the same formal sequence of actions and arrive at exactly the same specification independently. But, in practice, rarely do two designers arrive at exactly the same specification even when it is based on the same, simple set of requirements.

Differences arise because much is implied by a requirements description without being explicitly stated in it. Take, for example, the seemingly innocuous set of requirements presented to participants during the first experiment's English to Z translation exercise. From this, every participant managed to derive a different, but nevertheless consistent, specification of the same problem. This illustrates an important, but often overlooked, issue with regard to the production of formal specifications. These "implicit requirements" are normally implemented by designers according to their own discretions and personal styles of writing. In this case, participants appeared to make implicit but conscious decisions involving at least the following issues: the use of valid and invalid Z notation, the choice of meaningful identifier names, the data types assigned to each variable, the use of parentheses to clarify operator precedence, the ordering of expressions and the use of variables for storing intermediate results. Overall, the findings from this experimental task showed that the production of a formal specification is far from being a completely automated process, but is in fact frequently guided by informal actions and subjective human judgement. Hence, the engineering community should be careful not to underestimate the susceptibility of the formal specification process to human error.

Myth 6. *Using a formal notation constrains a writer's creativity.*

> "No doubt the logic is easy enough once one has studied it, but knowing how to construct an abstract model requires real understanding and experience."
>
> Oakley [Oak89].

The Sapir-Whorf theory of linguistic relativity [Who56, p. 27] argues that a person's thinking processes and behaviour is very much dependent upon the language in which thinking is conducted. The theory argues that the users of different grammars are pointed towards different kinds of observations, which eventually lead them to hold contrasting views of the world. Similarly, it is often hypothesised that formal notations constrain the way in which their users think and write because the grammatical constructs and rules that govern them are severely restricted in comparison with, say, those of natural languages. This can be shown to be a false belief because there is no evidence to suggest that people confine their reasoning to the deductive apparatus provided by formal systems. In fact, it is normally the case that, when people reason about a formal expression, they do so informally, using a mixture of natural and formal language. A typical thought process might be: "If the current value of variable s is *off*, then execution of the statement $s = \mathit{off} \Rightarrow s' = \mathit{on}$ would result in s' being *on* afterwards." So, whatever can be expressed in a formal language can be thought in a user's native, natural language. Furthermore, it must be realised that the Sapir-Whorf hypothesis was originally proposed as an explanation of the cultural differences that arose in different countries because of their use of different natural languages; it was never meant to be applied to compare the types of thinking evoked by technical and natural language grammars.

Responses from the first experiment's English to Z translation task failed to support the hypothesis that designers' thought processes are inhibited by their use of a formal notation. Although every participant offered responses corresponding to one of the two valid interpretations of its ambiguous English requirements, the fact that no two solutions were exactly the same proved that there are no formalised procedures for writing formal specifications. It also showed that designers must frequently employ their own subjective judgements, innovation, language expertise and experience in order to arrive at suitable solutions. For these reasons, it is the authors' opinion that the creativity in designers' problem-solving behaviour is not significantly impaired simply because they operate within the framework of a formal language rather than, say, a natural language. Furthermore, the task's findings suggest that formal notations, despite being more restricted than natural languages, are still sufficiently powerful to allow designers to exercise a large degree of creativity and freedom of expression.

Myth 7. *A quality specification is a verifiably correct specification.*

> "We do not argue that strict logical deduction should be the only way that mathematics should be done, or even that it should come first; rather, it should come last, after the theorems to be proved, and their proofs, are well understood."
>
> Maurer [Mau79].

Current research into formal methods appears to be progressing in two main directions: improving automated verification procedures and improving the readability of formal specifications. Although it might be a tremendous advantage for designers to be able to verify independently that selected properties of their specifications are both complete and consistent with regard to a client's requirements, it is debatable whether verified correctness should be the primary aim of software designers when one considers the role of a system specification in the overall development process. Since a software specification typically forms the basis from which future design or implementation work progresses, it is important that its readers are able to understand and reason about a specification accurately. In the past, imprecise or unintelligible specifications have led to developers making false assumptions or incorrect decisions which have had repercussions throughout the latter stages of software projects, causing the appearance of faults or anomalies in the system design or code produced. Although it might help to eliminate the number of previously undetected logical flaws and improve the verifier's understanding of a system, verification does not by itself increase the likelihood that the intended readers of a specification will be able to interpret more clearly and reason about it more effectively. This is because a specification might still be expressed in an unreadable or unnecessarily complex manner which could potentially stimulate erroneous human reason.

In order to communicate effectively with its audience, a specification must be readable and, in order to be readable, its designer must employ a style of writing that takes into account what is being specified and the intended audience's expertise. The findings from the first experiment are evidence of this. They illustrated how it is still possible to employ a range of contrasting writing styles using a restricted language and how even seemingly trivial requirements descriptions can still be implemented in a variety of subtly different ways. So, whilst it might be desirable for a specification to be verifiably correct with regard to properties of a customer's requirements, it is essential that it is easily comprehended by its intended audience, which might include programmers, designers, and managers. It is the authors' opinion that Gravell [Gra91] was closest to the truth when he said "clarity and comprehensibility are your main aims in writing a formal specification."

Conclusion

The initial study of the Human Cognition and Formal Methods research project has generated many results which appear to have implications for the future practice of software specification. It has uncovered many of the cognitive processes underlying the process of formal specification and served as a valuable learning experience. Overall, its results indicated that there still exist many misconceptions regarding formal methods and that the processes associated with formal specification may be more susceptible to human error than is generally believed. Its results also suggested that the ease with which a person can understand and reason about a formal specification is affected by various factors including its content and context, and personal characteristics of its readers and writers. Although the first investigation concentrated mainly upon deductive reasoning about conditional statements, its findings now provide a basis for future investigations which aim to test whether findings from past psychological studies of reasoning about negative, disjunctive and conjunctive statements also carry across into the domain of formal specification. It is hoped that these investigations will shed some light on the main sources of psychological complexity in understanding and reasoning about formal logic and formal specifications.

Acknowledgements

The authors would like to thank the following reviewers for their many helpful comments regarding this paper: Jean Baillie, Carol Britton, Ben Potter, Richard Potter, Richard Ralley, Jane Simpson, Mark Surman and Paul Taylor. The authors would like to thank Anthony Hall, Jonathan Bowen and Michael Hinchey for the inspiration of their original papers: "Seven myths of formal methods" [Hal90] and "Seven more myths of formal methods" [Bow94]. Finally, the authors would like to thank all of the staff and students who participated in the first experiment. The Human Cognition and Formal Methods research project is supported by Grant No. J00429434043 from the Economic and Social Research Council.[2]

References

[Bow94] J.P. Bowen and M.G. Hinchey, *Seven more myths of formal methods: dispelling industrial prejudices.* Oxford University Computing Laboratory Technical Report PRG-TR-7-94, Programming Research Group, Oxford, June 1994.

[Bra91] M.D.S. Braine and D.P. O'Brien, A theory of If: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, *98*, 182-203, 1991.

[Byr89] R.M.J. Byrne, Suppressing valid inferences with conditionals. *Cognition*, *31*, 61-83, 1989.

[DeM79] R. DeMillo, R. Lipton, and A. Perlis, Social processes and proofs of theorems and programs and programs. *Communications of the ACM*, *22*, 271-280, May 1979.

[Eva77] J.St.B.T. Evans, Linguistic factors in reasoning. *Quarterly Journal of Experimental Psychology*, *29*, 297-306, 1977.

[Eva83] J.St.B.T. Evans, Linguistic determinants of bias in conditional reasoning. *Quarterly Journal of Experimental Psychology*, *35A*, 635-644, 1983.

[Eva93] J.St.B.T. Evans, Bias and Rationality. K.I. Manktelow and D.E. Over (Eds.), *Rationality: Psychological and Philosophical Perspectives.* London: Routledge, 1993.

[2] This paper is based on a more comprehensive psychological report of the first experiment. Requests for this report or the experimental task sheet presented to participants should be sent to Rick Vinter, School of Information Sciences, Hatfield, Herts., AL10 9AB, UK. Email: R.Vinter@herts.ac.uk.

[Gra91] A. Gravell, What is a good formal specification? In J.E. Nicholls (Ed.), *Z User Workshop, Oxford 1990. Proceedings of the Fifth Annual Z User Meeting, Oxford, 17-18 December 1990*, Workshops in Computing, Springer-Verlag, 1991.

[Gri84] R.A. Griggs, and J.J. Chrostowski, Reasoning with realistic disjunctives. *Quarterly Journal of Experimental Psychology, 36A*, 611-627, 1984.

[Hal90] A. Hall, Seven Myths of Formal Methods. *IEEE Software*, *7* (5), 11-19, September 1990.

[Joh72] P.N. Johnson-Laird and J.M. Tridgell, When negation is easier than affirmation. *Quarterly Journal of Experimental Psychology*, *24*, 87-91, 1972.

[Kol73] P. Kolers, Translation and bilingualism. In G.A. Miller (Ed.), *Communication, Language and Meaning: Psychological Perspectives.* New York: Basic Books, 1973.

[Lak71] R. Lakoff, If's, and's, and but's about conjunction. In C.J. Fillmore and D.T. Langendoen (Eds.), *Studies in Linguistic Semantics.* New York: Holt, Rinehart and Winston, 1971.

[Lis79] B. Liskov and V. Berzins, An appraisal of program specifications. In P. Wegner (Ed.), *Research Directions in Software Technology*, Cambridge, Mass: MIT Press, 1979.

[Loo94] M. Loomes, D. Ridley and D. Kornbrot, Cognitive and Organisational Aspects of Design. In F. Redmill and T. Anderson (Eds.), *Proceedings of the Second Safety-Critical Systems Symposium, Birmingham UK. 8th to 10th February, 1994.*, 186-193, Springer-Verlag, 1994.

[Mau79] W.D. Maurer, Letter to the editor. *Communications of the ACM*, *22*, 625-629, November 1979.

[Oak89] B. Oakley, Opening address: The state of use of formal methods. In J.E. Nicholls (Ed.), *Z User Workshop: Proceedings of the fourth annual Z user meeting, Oxford, 15 December 1989*, Workshops in Computing, Springer-Verlag, 1990.

[Pot91] B. Potter, J. Sinclair, and D. Till, *An Introduction to Formal Specification and Z.* Hemel Hempstead: Prentice-Hall, 1991.

[Spi92] J.M. Spivey, *The Z notation: a reference manual.* Second Edition. Prentice Hall International, 1992.

[Was66] P.C. Wason, Reasoning. In B.M. Foss (Ed.), *New Horizons in Psychology. Volume 1*, Reading: Penguin, 1966.

[Was72] P.C. Wason and P.N. Johnson-Laird, *Psychology of Reasoning: Structure and Content.* London: Batsford, 1972.

[Who56] B.L. Whorf, *Language thought and reality: Selected writings of Benjamin Lee Whorf.* Cambridge, Mass: MIT Press, 1956.