MANON PENNING

Biased Toddlers

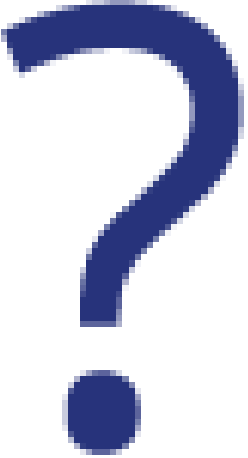How bias will affect even the simplest of specifications

"Biased? No, as a requirements engineer I strive to be impartial and unbiased", right? But really being impartial and unbiased is extremely difficult! That is, if it is at all possible. People are influenced by their current belief system, their mental image or models of the world. Those in turn are shaped throughout our lives based on the cultures with which we identify ourselves, our background in terms of education and training, and life's experiences.

We use these models and derived heuristics to take shortcuts because considering all possible options in an exploratory thought experiment rarely is an option. This phenomenon closely relates to a certain form of cognitive bias: confirmation bias. It can easily be illustrated using an advanced variant of a game called 'The Genius Toddler Challenge' described in RE Magazine 2016-01 (Penning, 2016). The effects were way more prominent than anticipated.

When somebody uses an event that has occurred as confirmation of a vague statement in a horoscope, or when a person seems to only register the good things about another person they are madly in love with, confirmation bias is in play. Confirmation bias is explained slightly differently by various psychologists, but in this article the broader application of the term *"searching for evidence, interpreting it, or recalling it from memory"* is used [5]. Specifically applied to 'The Genius Toddler Challenge' this comes down to bias in searching for information (and which parts of and how to capture that information), and bias in interpretation.
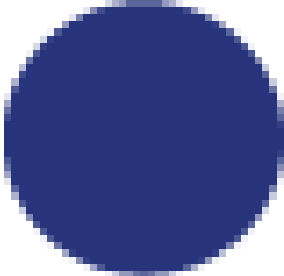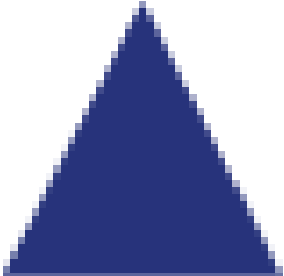
The basics

In 'The Genius Toddler Challenge' [4] the basic concept behind a simple game to illustrate some of the difficulties requirements engineers face was introduced. For this challenge, participants are divided into teams of three to six people and each team is given the assignment to write down a description of two abstract images using only language a toddler could use. Subsequently the descriptions of each team are distributed to two or more of the other teams, who are asked to reconstruct the original image. These assignments are a lot harder than they look and provide ample food for discussion.

The images provided can be any set of simple drawings made using a basic drawing tool.

For the original Genius Toddler Challenge images were used consisting of 3 lines. These lines are equal in dimensions, differently coloured. Each drawing has a unique arrangement of the lines.

In the remainder of this article the concept of 'The Genius Toddler Challenge' and its two phases are assumed to be already understood.

Introducing bias

By adding one extra level to the original challenge, bias can be provoked. Instead of using images consisting of instances of the exact same basic shape, the advanced version uses a different basic shape per team as illustrated below for four teams.

Team A	Team B	Team C	Team D
——————	●	■	▲

In addition, filling can be introduced. Team A might get all solid fills, Team B only borders and Team C shaded shapes or a mixture of the first two unless a Team D is present. The team with shaded shapes can get an extra challenge when the direction of shading is significant, as illustrated below.

Differently shaded squares
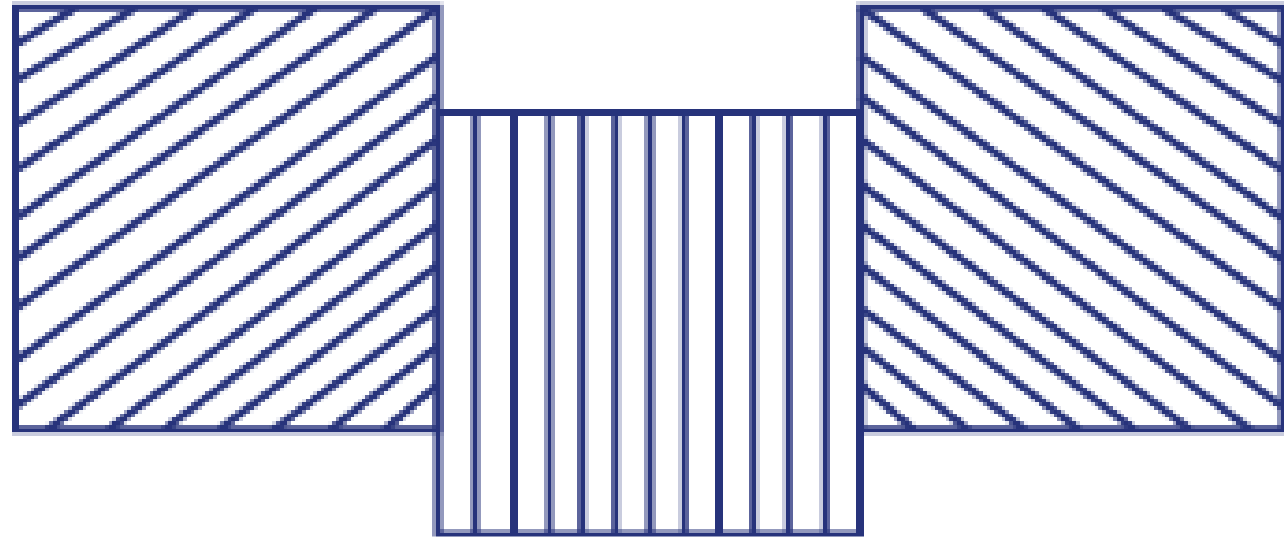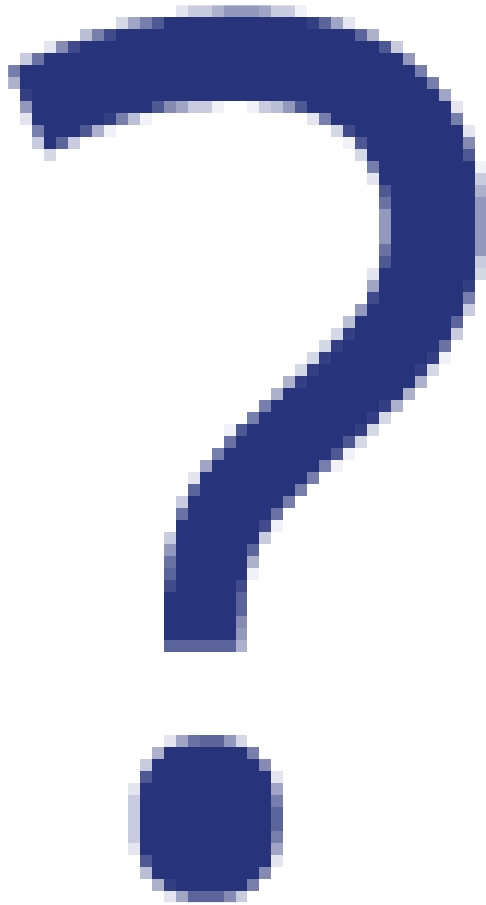
In the introduction to the challenge, participants are not explicitly made aware of the different shapes and fillings. The facilitator will simply tell them that they will receive similar abstract images made up of a basic shape. No explicit example is shown during the introduction because it might provide clues to teams with a different basic shape or filling to that shown in the example.

Apart from changes to the images provided and a slightly more mysterious introduction, no changes are made to the original challenge. The two phases of the challenge remain intact. In fact, when trying to illustrate bias it is all the more important that in Phase 2 each description is handed out to multiple teams. If only one team were to draw each description, the differences in bias will not become as apparent as they might when two or more teams recreate the same image. Alternatively, the facilitator can point out bias in the descriptions themselves such as the absence of relevant information about the shape or filling. But pictures drawn from the same description by different teams might convey more than a lot of talking by the facilitator.

An example

In 2016 the advanced challenge described in this article was held a couple of times. One of those instances was with a company employing people with a focus on testing. Traditionally testers have often been perceived as that annoying bunch who complain about poor quality requirements – with some exceptions of course! Testers have always been stakeholders of requirements but too often they were included late. Nowadays, especially in agile contexts, the door is wide open for testers to contribute to the specifications early on. When using 'Specification by Example' (SbE) for instance [1], testers can be very helpful in providing sharp key examples. A possible pitfall of teams using SbE and testers in particular is a tendency to add too many examples that are not necessarily proper key examples in their own right. But that is a can of worms to be discussed elsewhere, just like whether SbE should be called 'Specification AND Examples', as examples alone do not a specification make [2].

When going through the challenge with that group of testers, a couple of things happened. First of all, they were surprisingly fanatical and unbelievably precise in the things they did add to their description. Because use of a rule was forbidden they started measuring in iPhone units only to discover that maybe using fractions was not such a bright idea. This spiralled into a discussion about types of iPhones and their differences in size and on to other makes and models of phones running Android, Microsoft, Ubuntu, etcetera. Other teams used their hands as measure and introduced thumbs and palms, or measured by the pen provided. One of the images provided to Team A looked like the image shown below. Next to the image a couple of key points from the description written by Team A during Phase 1 are listed.

| Original image Team A	Key points from the description written by Team A:
	• Used 'narrow rectangles' as description for the basic shape
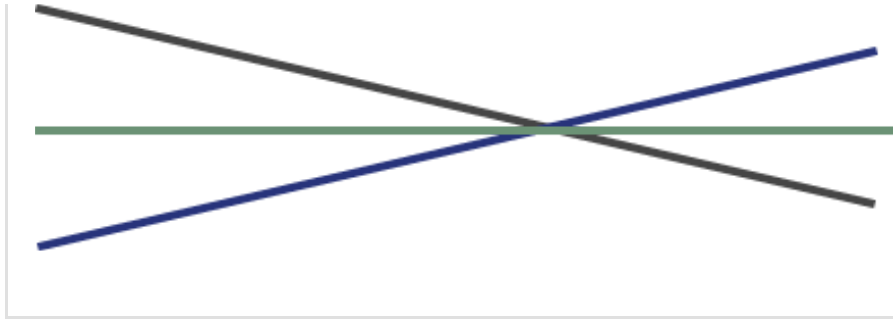
- Forgot to make the solid filling explicit ('one green, one black, one blue')

- Used 'wigwam' as a metaphor

Teams B and C (there were only three teams in this case) recreated the image in Phase 2. The basic shapes those teams were familiar with are as shown in the example earlier with four teams (circles for Team B, squares for Team C), but with no filling and shading respectively.
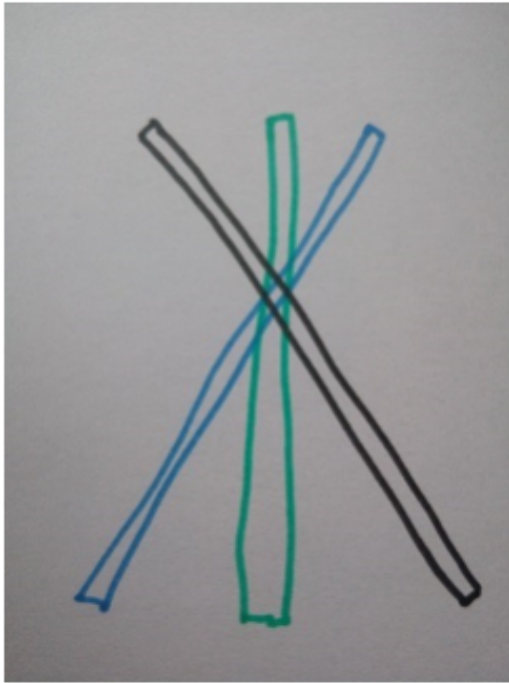
Team B
(used to circle, no filling)
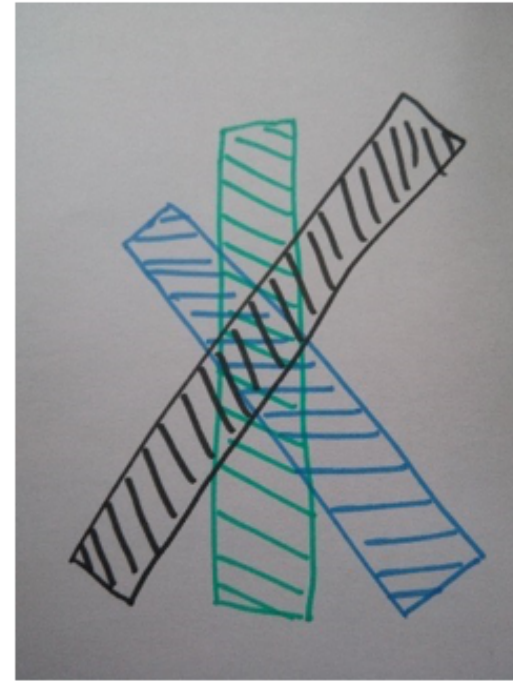
Team C
(used to squares with shading)

Need I say more?

Notice that for some reason Team B constructed a mirror-image. The basic shape drawn by Team B more closely resembles the original image, possibly because one of the team members had participated in the original challenge a year or more back. In both drawings it is indistinguishable which rectangle is on top. Our brains seem to tell us the black one is on top but in fact in both cases the green one was drawn last. Only once was a

team encountered that left out parts of the lower rectangles to make the stacking order clear. That team had two graphical designers in their midst – maybe that had something to do with it.

Evaluation and wrap-up

The original version of The Genius Toddler Challenge discussed issues related to the typical symptoms of inadequate RE, and typical reasons for this ([3] , p. 8) were pointed out. With the more advanced version as described here, these are still likely to appear. Additionally the missing requirements are shown to be closely tied to bias. When writing down their specifications in Phase 1 the teams will use their current model to describe the images provided. That is likely to lead to missing requirements because the teams tend to leave out certain details that may be highly relevant, such as the type of filling. In Phase 2, when drawing the images based on specifications from another team, teams will assume certain details are equivalent to what they have seen in their images in Phase 1 when given the room for interpretation.

Of course this challenge is a gross over-simplification of what will be encountered in the real world. The lesson to take home for people dealing with specifications, both writing and interpreting them, is to be aware of effects that bias can have and to (continue to) strive to minimize the impact on the end product. On the other hand, checking and researching everything in minute detail is impossible, so it would be wise to fully research high-risk items. But assumptions will have to be made and heuristics applied. Be aware of the associated pitfalls and weaknesses!

For more information regarding the (Biased) Genius Toddler challenge, for instance to be provided with the full set of original images or to provide information about a fully facilitated challenge, please send an email to: info@improveqs.nl.

References and Literature

[1] Adzic, G. (2011). Specification by example: How successful teams deliver the right software. Manning.

[2] Gerrard, P. (2012). Specification by Example is not Enough. Retrieved from Paul Gerrard's blog: http://www.gerrardconsulting.com/?q=node/596

[3] IREB. (2015, March 1). Syllabus IREB Certified Professional for Requirements Engineering Foundation Level.

[4] Penning, M. (2016). The Genius Toddler Challenge. Requirements Engineering Magazine (2016-01: Seeking new perspectives). Retrieved from https://re-magazine.ireb.org/

[5] Risen, J., & Gilovich, T. (2007). Informal Logical Fallacies. In R. J. Sternberg, H. L. Roediger III, & D. F. Halpern, Critical Thinking in Psychology (pp. 110–30). Cambridge University Press.

Manon Penning

Manon Penning has a broad background in the fields of IT and education. She is an educational specialist, requirements engineer, quality analyst and trainer currently employed by Improve Quality Services, a provider of consultancy and training. Via Improve Quality Services she has been working for several Dutch companies over the past few years.