How do low educated audiences understand motion in static visuals?

Two explorative studies into low educated South Africans’ comprehension of motion suggestion in static visuals

Hanneke Hoogwegt
December 2007
Tilburg University

1st Supervisor: Prof. dr. Fons Maes
2nd Supervisor: Prof. dr. Emiel Krahmer
Preface and acknowledgments

Finally, in February 2006 Karen and I sat in the airplane, flying towards our South African adventure we had been looking forward to for so long. And an adventure it certainly was. We went through quite some unforgettable experiences in one of the world’s most beautiful countries. From spotting lions, hippos and giraffes in Kruger National Park to partying with our fellow international students from all over the world (we were still students after all…). Of course, our main goal was to collect the research data for our MA theses, which we – after quite some stress attacks by the undersigned and a lot of networking efforts – successfully did. We ignored all warnings and visited one of Pretoria’s townships with (not even our own) laptops, microphones and webcams until we had reached our targets. Karen, thanks for being part of this wonderful experience, we were (are) a good team!

Today, it has been almost two years since I left for South Africa. In between the courses of the researchmaster, I worked hard to finish my thesis. It took a bit longer than expected, but with the help of a number of persons it is finally finished now. My first credits go out to my supervisor Fons Maes for giving me the chance to participate in the Epidasa Project and getting in touch with South Africa and the South African people. It has been a blessing to devote a great amount of my researchmaster hours to such a fantastic project. I also appreciate all the help and the pleasant cooperation throughout the whole process and your patience when things went not so well. In South Africa, Adelia Carstens had been of great help in achieving our targets there, so thank you Adelia for all your efforts. I also appreciate the efforts by the staff members of the Center for the Study of AIDS, who were of great help with finding interpreters, respondents and a location for the interviews. Furthermore, I would like to thank Carel van Wijk for the great help with the statistical parts and Lennart van de Laar for his help with the research materials.

I also owe a big thank you to my parents, for supporting me in every imaginable way and believing in me. Who also deserve to be mentioned are my sisters Karlijn and Annelieke. My friends (especially the Chickieś) deserve a big thank you as well, for all the fun during our student days!
Last but certainly not least: thank you Pim, for your love and your patience every time I was freaking out or when I couldn’t find time for you because I had to pass all my courses with at least an 8 or was too busy with organizing all kinds of things at the university.

I really enjoyed my student days. On the one hand I feel sad because those days are over now, but on the other hand I look forward to applying my acquired knowledge and skills in practice (and I will make sure that the cartoon below will never apply to me…).

Oirschot, December 18th, 2007

Hanneke Hoogwegt
Abstract

This thesis reports on a research project aimed at gaining insight into how motion in static visuals can be effectively suggested for low educated South African audiences. In Study 1, structured interviews about visuals displaying uncomplicated actions were held with fifty township inhabitants, who were all speakers of African languages. The responses were coded and analyzed both qualitatively and quantitatively. Consistent with previous research in a Western context, the depiction of the arrow, a commonly used abstract visual element to indicate (specific) motion(s), was found to trigger motion and direction of motion. The amount of visuals with arrows spontaneously described in terms of motion was found to increase as the education level of the respondents increased, implying that the relatively higher educated respondents benefit most from the presence of arrows. This finding is consistent with previous research into the visual literacy of the population of the current study, which already revealed that higher educated viewers tend to be more familiar with conventional representations of abstract notions than lower educated viewers. However, as can also be predicted by previous research into the visual literacy of low educated South Africans, a substantial amount of the respondents was not familiar with the meaning of the arrow as an indicator of motion and direction of motion. Furthermore, the results obtained in Study 1 indicate that in addition to this abstract visual element, the analogical element ‘human hand’ (which in contrast to the arrow was familiar to all respondents) also has the expressive power to trigger motion and direction of motion. This finding gave rise to Study 2, which aimed at gaining additional insights into the expressive power of hands. The obtained results reveal that for visuals intended to display the movement of (an) object(s), displaying hands in these depictions triggers motion and direction of motion to a similar or even greater extent than displaying arrows. Furthermore, the combination of hands and arrows was found to trigger far more responses in terms of (the intended) motion compared to when only an arrow is displayed. It appears that hands suggest the presence of an actor, which again suggests the presence of (a specific) motion. This finding is consistent with theories on and experimental research into embodiment and affordances, in which it is argued that humans assign meaning to what they perceive thought their own (bodily) experiences. For designers, these results imply that arrows can be used to indicate motion and direction of motion in visuals aimed at low educated South African audiences, but that it should be taken into account that not all viewers may be familiar with the meaning of this abstract element in the particular context and that the expressive power of familiar, natural elements such as hands, should be exploited.
# Contents

1. INTRODUCTION 7

2. STUDY 1: SUGGESTING MOTION IN STATIC VISUALS 9
   WITH AN ABSTRACT VISUAL ELEMENT (THE ARROW)
   2.1 The power of visuals 9
   2.2 Are pictures always worth a thousand words? 11
      2.2.1 Interpretation of visuals in the (South) African context 12
      2.2.2 The influence of education level on the interpretation of static visuals 16
      2.2.3 Summary 18
   2.3 The arrow as an effective abstract visual element for suggesting motion in static visuals 18
   2.4 Research questions and hypotheses 21
      2.4.1 Main research questions and hypotheses 21
      2.4.2 Additional research question and hypotheses 22
   2.5 Method 23
      2.5.1 Respondents 23
      2.5.2 Research materials 24
      2.5.3 Interview setting 26
      2.5.4 Procedure 26
      2.5.5 Design 27
      2.5.6 Analyses 27
   2.6 Results 29
      2.6.1 Effect of arrow on expression type 29
      2.6.2 Interpretation of arrow 31
      2.6.3 Effect of education level and age on expression type 31
      2.6.4 Sources for correct dynamic expressions 33
   2.7 Conclusions and discussion 35
      2.7.1 Effect and interpretation of the arrow 35
      2.7.2 Effect of education level and age 37
      2.7.3 Other visual elements suggesting motion 38

3. STUDY 2: SUGGESTION MOTION BY ANALOGY (HANDS) 40
   3.1 Introduction 40
   3.2 Research questions and hypotheses 40
   3.3 Method 42
      3.3.1 Respondents 42
      3.3.2 Research materials 42
      3.3.3 Interview setting 43
      3.3.4 Procedure 43
      3.3.5 Design 43
      3.3.6 Analyses 45
   3.4 Results 44
      3.4.1 Hands versus arrows 45
3.4.2 Hand and arrow versus arrow only 46
3.4.3 Interpretation of arrow 48
3.4.4 Sources for correct dynamic expressions 49
3.4.5 Effect of education level and age on expression type 50

3.5. Conclusions and discussion 51
3.5.1 Hands instead of arrows 51
3.5.2 Hands in addition to arrows 51
3.5.3 Interpretation of arrow 52
3.5.4 Effect of education level and age 53

4. GENERAL DISCUSSION 54
4.1. General discussion and conclusions 54
4.2. Limitations of the research 58
4.3. Suggestions for future research 60

BIBLIOGRAPHY 63

APPENDICES
1. Introduction

Who has ever visited South Africa must have noticed the many billboards in the landscape, such as the ones below, part of a LoveLife\(^1\) campaign.

These billboards are not there without a reason. Almost 19% of the South Africans are infected with HIV/AIDS. Furthermore, estimations indicate that about 29% of the pregnant women are infected with the virus (UNAIDS/WHO Epidemic Update, 2007). As a consequence, there is a great need among the South African population for information about HIV/AIDS, but also about other health related topics. Obviously, it is important that the messages in education and intervention campaigns can be understood by all readers and viewers. The use of mass media, such as the billboards in figure 1, pamphlets and information brochures is an attractive option in terms of finance, time and reach. Major intervention organisations such as Soul City\(^2\) and LoveLife have made extensive use of such printed materials, to which many South Africans are exposed. However, effective communication about (sometimes complex) health related messages to South African audiences by non personalised mass media is quite a challenge. The population is very diverse in terms of for example ethnicity, first languages spoken, education level and literacy level\(^3\). In many of the information materials, visual messages – often in addition to text – can be found. At first sight visuals indeed seem a valuable source of information, especially in a country where the population is so diverse and literacy and education levels are relatively low among a substantial amount of the population. To illustrate the latter: 30% of the South African population is illiterate and an other 40% has limited reading skills (Carstens, 2004a). Moreover, as Carstens & Snyman (2003) concluded, basic instructional materials on health issues have a readability level comparable to Grade 9. Intuitively, the use of visual messages instead of or in addition to textual messages might seem a plausible and effective solution

\(^{1}\) LoveLife is an organisation which sets up intervention programmes in South Africa aimed at pre- and newly-sexually active adolescents (ages 12 – 17)

\(^{2}\) Soul City is a South African institute for health and development communication that among other things spreads print materials about HIV/AIDS and other health related topics among the South African population

\(^{3}\) Eleven official language communities are distinguished in South Africa, as well as four population groups (African, coloured, Indian/Asian and white); education levels vary from no formal schooling (7%, especially older South Africans) to a degree or higher (3.7%) (AMPS, 2005)
for informing diverse populations, including persons with limited reading skills. After all, the proverb *a picture is worth a thousand words* does not seem to exist by coincidence. However, even though the messages in for instance the billboards in figure 1 are partly explained by visual information, they require quite some verbal and visual literacy skills. The viewer is for example required to understand the fusion of the condom and the lips or what it means that “HIV loves skin on skin”, while at the same time the visual message suggests only one body.

A look into public information documents spread in South Africa reveals that designers often intend to express a scene or action including motion. Some visuals intend to depict complete scenes and situations, as is the case for figure 2. This visual depicts a scene in which a female is trying to keep away the three children from a bleeding child (in order to prevent them from becoming HIV-infected by blood contact, making the visual of an advisory nature).

Many printed materials also include visuals designed to instruct the viewer, for instance regarding how to put on a condom or how to use an asthma pipe (figure 3). A precondition for comprehending the instruction, is that the viewer recognizes the intended action and corresponding motions. With respect to figure 3, the viewer is expected to recognize the pressing movement of the upper finger, part of the action of using an asthma pipe as it should be. Such a rich meaning is in sharp contrast with the static nature of visuals. In printed communication materials such as brochures, posters and billboards, the designer can only try to *suggest* motion (as is done in figure 3 by means of an abstract symbol: an arrow). It can not be directly visualised. At the same time, visual communication about for instance health related topics often requires the depiction of *dynamic* actions such as the ones displayed in figure 2 and 3. The focus of the present study is on how motion can be effectively suggested in static visuals for low educated South African audiences. In this thesis the reports of two studies into the comprehension of motion in static visuals by low educated South Africans are presented. The first study focuses on the possible effect of an abstract visual element (the arrow) on the interpretation of static visuals. This study gave rise to a follow-up study, in which the focus lies on suggesting motion by analogy: the human hand.
The first study is presented in the following chapter (chapter 2), the second study is presented in chapter 3.

2. Study 1: suggesting motion in static visuals by an abstract visual element (the arrow)

The theoretical framework of Study 1 (consisting of section 2.1, 2.2 and 2.3) provides an overview of literature on the interpretation of visuals in the (South) African context in general, as well as research into suggesting motion in static visuals by an abstract element (the arrow). First of all, the power of visuals revealed by various studies will be outlined in section 2.1 These results indicate that visuals can be powerful communicative means. However, research in the (South) African context, discussed in section 2.2, shows that this is not always the case and that pictures do not always seem to be worth a thousand words. For instance the visual representation of abstract notions which cannot be visualised directly, seems quite a challenge. One of these abstract notions which cannot be directly visualized by static visuals, is the one of central focus in this study: motion. Section 2.3 deals with the effect of a stereotype method for suggesting motion by an abstract visual element: the use of arrows. The literature discussed in section 2.1, 2.2 and 2.3 leads to various research question and hypotheses, which will be discussed in section 2.4. The method of the study will be outlined in section 2.5, followed by an overview of the results in section 2.6. Conclusions will be drawn from these results, and they will be briefly discussed in the final section of this chapter (section 2.7).

2.1 The power of visuals

The power of visuals has been emphasised by various experts. According to Maes & Schilperoord (2002: 158,159), visuals can be included in a message in order to exert a positive influence on perception, decoding, comprehension, integration, transfer, attitude and intention. They argue that visuals have the ability to attract attention, to facilitate the (mental) process of decoding a message and enriching its meaning, and that visuals can contribute to attitude change, for example by activating emotions and associations. This is in line with various scholars who have also emphasised the power of visuals when it comes to increasing the speed of message transfer, attracting attention, stimulating motivation, improving comprehension and enhancing recall (e.g. Glenberg, 1997, 2002; Mayer, 1999; Mayer en Gallini, 1990; Mayer & Sims, 1994; Paivio, 1986; Sinatra, 1986). However, it must not be left unmentioned that these studies do not focus on visuals as stand alone, but mainly on combinations of modalities (spoken text, written text, printed visuals, animated visuals). This is also the case for the positive influence of visuals found in the field of health communication, where Nogh and Shepherds
(1997b) found culturally appropriate pictorial medication schemes accompanying spoken instructions to be much more effective in terms of comprehension and compliance than spoken text alone. Houts et al. (Houts et al. 1998; Houts et al., 2001) found the recall of medical instruction to increase enormously when pictorial reminders were added to instructions. Houts et al. (2006) reviewed studies in the fields of health education, psychology, education and marketing and came to similar conclusions. According to them, “visuals closely linked to written or spoken text can, when compared to text alone, markedly increase attention to and recall of health education information.” Furthermore, the studies show that comprehension can be improved when visuals are showing relationships among ideas, or spatial relationships. Moreover, visuals were found to have the ability to change adherence to health instructions. Houts et al. (2006) conclude that all patients, but in particular low literates, can benefit from the use of visuals in health education materials. It must be noted that also in these studies, the visual messages were part of a broader intervention, also including for instance spoken or written text.

Though the methods used for the data collection are not always clear, several studies conducted in the (South) African context reveal positive effects of visuals as well. During his research into the use of printed materials by small-scale farmers in South Africa’s rural province KwaZulu-Natal, Stefano (2004: 60) found visuals to be positively judged by the viewers. Their comments “When I see the visual I become interested in the story and read about it”, “Pictures are good” and “It is easy to understand the pictures” illustrate the positive attitude towards the use of visuals in printed materials. A study by Carter (1999: 65) into rural farmers’ use of agricultural extension materials in Uganda and Ghana provides additional support for the use of visuals. The farmers preferred designs with plenty of clear illustrations and little text. Dowse & Ehlers (2004) conducted a study into the comprehension of medicine instructions among 304 low-literate South African respondents from eight out of South Africa’s nine African language groups. They found almost all respondents (98%) to respond positively to the idea of having pictograms in their medicine labels.

Can we conclude now that visuals are perfect communicative means, solving all obstacles raised by the diversity of the South African population in terms of e.g. language, culture, education level and literacy level? Probably not. The positive findings by the studies conducted so far in the (South) African context mostly relate to the attractiveness of visuals, leaving the comprehension aspect unmentioned. With respect to comprehension, among others Hoffmann (2002) warns for an overestimation of the power of visuals. It is assumed that printed media with visuals provide a language-independent, supracultural means of communication (Work, 1990; Hoffmann, 2000: iv), while environment, cultural base, education level, values and communication influence each individual’s perception and interpretation of visual information (Carstens et al., 2006; Doak et al., 1996; Dowse & Ehlers, 2004; Foesenek, 2006; Moynihan & Narayanann, 1981; Zimmerman & Perkin, 1982). The reliance on visuals may cause problems when their true potential is underexploited and overestimated, as research into the interpretation of visuals by for example low literate and low
educated (South) African target audiences has revealed (e.g. Hoffmann, 2002; Dowse & Ehlers, 2004; Carstens et al., 2006; Foesenek, 2006). In the section 2.2, an overview of this research is presented.

### 2.2 Are pictures always worth a thousand words?

The first reports about difficulties experienced with visual illustrations and script, came from missionaries in the 70’s. Before, it had been taken for granted that visuals have the ability to communicate meaning across the boundaries of language and culture (Hoffmann, 2002). According to Hoffmann (2002), this “naïve expectation of the universal understandability of pictures” lasted until the late 1970s. Hoffmann (2002) provides an overview of many explorative field studies into visual literacy in the African continent. Unfortunately, much of this research is dated. In addition, most of the data are collected by small-scale participatory field studies, making generalization and reliability highly questionable. Moreover, as Hoffmann (2002: 136) remarks: “almost all studies lack a purposeful theoretical orientation”. Despite these limitations the overview is valuable, as it provides clear indications that visual messages often include pictorial conventions, of which some have become so natural to skilled viewers and designers that the conventional nature of certain elements is easily overlooked. This can be illustrated by the following example in the overview by Hoffmann (2002: 142).

Many Western oriented viewers might experience difficulties in finding an element of a conventional nature in figure 4. In Western illustration there is an artistic tradition to use highlights (e.g. white marks) for suggesting reflections on shiny surfaces. For the ones familiar with this pictorial convention, the use of highlights (white marks) in the eyes of the boy in figure 4 is an easily understood indication of sparkling eyes. However, inhabitants of a Kenyan township considered the eye to be damaged, and the teeth as grown together. They seemed to interpret the visual ‘literally’, resulting in undesired responses, given that the visual was part of a toothpaste advertisement...

Various dated and several more recent studies into visual literacy in the (South) African context have revealed insights in line with the findings of the field studies referred to by Hoffmann (2002). These studies, and the influencing factors (e.g. familiarity with the depicted elements and literacy/education level) with regard to the interpretation of visuals they reveal, are discussed in the following section.

---

4 The study by Foesenek (2006) is a replication study of the one by Carstens et al. (2006), to improve several limitations of the study by Carstens et al. (2006: 229)
2.2.1 Interpretation of visuals in the (South) African context

In order to understand a visual’s meaning, a successful interpretation of various aspects by the viewer is required. This can be quite a challenge, given that “static visuals highly rely on implicit meaning and suggestion provided by a complex interplay of natural (analogical), iconic (associative) and symbolic (conventional) elements” (Maes et al., in press). In the following three sections, results obtained by studies conducted in the (South) African context into the interpretation of the various visual elements and pictorial conventions, is outlined.

Interpretation of natural (analogical) visual elements

Several studies (e.g. Foesenek, 2006; Carstens et al., 2006) among South Africans have shown that natural visual elements such as objects and humans familiar to the viewer, are easily recognized. To understand the meaning of a visual, being familiar with among others the natural, analogical elements depicted in the visual, is essential (by ‘natural’ and ‘analogue elements’ is it referred to concrete, perceivable elements in reality such as humans and objects). This sounds very obvious, but at this point wrong assumptions are easily made by senders or designers of communicative messages. The latter is for instance revealed in the study by Dowse & Ehlers (2004) into the comprehension of medicine instructions among 304 low-literate South Africans. Their aim was to investigate “the influence of collaboration with the target culture on the design and interpretation of pharmaceutical pictograms by comparing the interpretation of symbols from the USP-DI with visuals designed locally by [the] research group, and to evaluate the interpretation of these pictograms by low-literate subjects from various South African language groups”. Overall, the locally designed instructions were preferred by the respondents. Moreover, they were much better understood compared to the standard ones by the USP-DI. This can be partly explained by the unfamiliarity of the participants with natural elements depicted in the USP-DI instructions. To illustrate this, Dowse & Ehlers (2004) found that a tablet box was hardly recognized as such, because in South Africa patients commonly receive their tablets in plastic bags and not in a box. In Zimbabwe, Cornwall (1992) studied the effectiveness of standard visuals and diagrams of reproductive anatomy, which are shown in clinics all over the world. She found the visuals to be ineffective in Zimbabwe, due to the fact that they represented the Western medical model of reproductive anatomy and the way the female reproductive system works. Visuals reflecting ‘ordinary’ people’s knowledge of the body, were interpreted more successfully. These visuals were based on body maps and diagrams drawn by local women on paper or on the ground. The locally developed drawings reflected knowledge from more practical experiences about the

---

5 USP-DI is short for United States Pharmacopeia Dispensing Information. “USP Pictograms are standardized graphic visuals that help convey medication instructions, precautions, and/or warnings to patients and consumers. Pictograms are particularly helpful in passing on important information to patients with a lower level reading ability and patients for whom English is a second language” (www.usp.org)
working of the body, as well as experiences related to pregnancies and advice from peers and older women.

The results obtained in the studies referred to in this section, stress the importance of taking into account what the target group is familiar with, i.e. what they are exposed to and therefore recognize from their own observations and (bodily) experiences.

**Interpretation of symbolic (conventional) and iconic (associative) elements representing abstract notions**

Natural elements referred to in the previous section are just a subset of a wider range of elements that need to be interpreted as intended in order to understand a visual as a whole, as well as its function. Verbal language provides the ability to refer directly to abstract notions which cannot be perceived, such as the notions of ‘love’ and ‘HIV/AIDS’, as well as descriptions of movement, action, figure ground, personal relations or the intention behind a message. Because these notions do not have an analogical residue in reality, they cannot be visualised directly. As a consequence, a wide range of visual elements and pictorial conventions have been invented to represent or suggest abstract notions. Examples of elements representing abstract notions can be found in the figure 5, 6, 7, and 8 (all used in South African health communication documents).

These elements show that visual elements representing abstract notions may vary in the extent to which they are purely conventional or not, i.e. the range from purely conventional to purely natural, analogical elements is of a gradual nature. Some of the elements have a purely arbitrary, symbolic, conventional relation with their referent (e.g. the mathematical symbols) while others have a more iconic, associative relation with the referent (e.g. the soda and the white bread, they are a part of a greater whole which is intended to be suggested by this part). Familiarity with these symbolic and iconic elements and the way they are supposed to be interpreted is required in order to assign the intended meaning to these elements. However, especially for low educated viewers, chances are that they are unfamiliar with (some of) these elements. This is for instance revealed in the studies by Carstens et al. (2006), Foesenek (2006), in which e.g. figure 5, 6 and 7 were included. One of the
findings by Foesenek (2006) was that the heart as a conventional symbol for love was not familiar to many low educated respondents, and that monster representing HIV/AIDS was often interpreted literally. Carstens et al. (2006) found the proportion of visual elements recognized to decrease as the elements were less familiar and less related to the respondent’s own (bodily) experiences.

Similar insights were revealed in the study by Dowse & Ehlers (2004) (in which e.g. figure 8 was included). They for instance found the R-convention to be poorly understood. Furthermore, they found visual metonymy to cause interpretation problems. Fire representing heat was not always recognized as such. One of the instructional visuals included a sun indicating that the depicted medicine should be taken during the daytime and a moon indicating that it should be taken at night. Several participants responded that the medicine should be taken when the sun or the moon can be seen and that on a cloudy day they would therefore not have to take the medicine.

The studies and examples outlined above illustrate that – even when the relation between the visual element and its referent is not purely conventional – not all viewers can be expected to be familiar with all (partly) conventional visual elements representing abstract notions, and/or the pictorial conventions required for a correct interpretation. This is especially the case for low educated viewers (see also section 2.2.2).

Comprehension of visuals beyond the level of separate (natural, iconic or symbolic) visual elements

So far, obstacles for the interpretation of visuals relating to separate visual elements were outlined. However, visuals are of holistic nature and in order to understand the message of a visual, a richer interpretation than just the identification of the elements (e.g. ‘I see a man and a woman’ for figure 5) is expected from the viewer. On the basis of the separate elements and their relations, a visual also has to be understood in terms of a state of affairs, of an action or a certain event, which can be expressed by a proposition (e.g. ‘a man and a woman who are expecting a baby and are worried about their baby being HIV-positive’ for figure 5). The visuals in the study by Carstens et al. (2006) and Foesenek (2006) are supposed to be interpreted in terms of an advice or recommendation relating to HIV/AIDS. Respondents need to translate a neutral state of affairs depicted in a visual into a recommendation (e.g. ‘when an HIV-infected woman is pregnant, it should be tested whether the baby is also infected with HIV’ for figure 5), often without any visual cues triggering such an interpretation. What is done in a visual (e.g. ‘a mother talking with her daughter about how children come to the world’) is supposed to be translated into what one should do (e.g. ‘as a parent I should talk with my children about how children come to the world’). Carstens et al. (2006) and Foesenek (2006) found the recognition of the basic intention of the visuals to differ among visuals and education levels. Recommendations consisting of a relatively complex message were, especially by the low educated respondents, often not recognized. The unfamiliarity of respondents with various iconic and symbolic elements which were poorly recognized, are likely to have contributed to these results. The symbolic and iconic elements are after all essential to recognize the state of affairs or event displayed by the visual. This recognition is again a precondition for understanding the intention behind a message, for familiarity with symbolic or iconic elements can be required as well. According to
Carstens et al. (2006) their data suggest that – especially for low educated respondents – “the actions and mental states of human beings can be more successfully expressed by using bodily expressions that are familiar on the basis of daily experiences than by applying literacy conventions”. The act of speaking for example was far more successfully indicated by a corresponding facial expression than by speech balloons. This once again demonstrates the power of recognizable familiar elements for the viewer, as opposed to purely conventional elements for which one is dependent on what one has learned instead of one’s own experiences.

Figure 9 is an example of a depiction that appeals to familiarity with pictorial conventions of yet a different nature. The viewer is expected to be familiar with the phenomenon of reading from left to right, in order to interpret the visual as a depiction of four separate actions of taking a medicine, representing different moments in time, chronologically depicted from left to right. This left-to-right interpretation required might be very natural to literate viewers who read from left to right, but Dowse & Ehlers (2004) found their non educated South African respondents to be unfamiliar with this convention.

Another, seemingly natural convention is the depiction (or: suggestion) of figure-ground. How elements in the visual relate to each other in terms of foreground and background is often pictorially expressed by the manipulation of space and size. The depiction of element X in the lower part of the visual and element Y in the upper part of the visual, may indicate that – from the perspective of the viewer – element Y is supposed to be interpreted as spatially further away than element X. This effect is strengthened by the manipulation of size: element Y being depicted relatively small compared to element X, once again indicates the distance between the two elements and the relative distance of both elements from the viewer. Figure 6 shows such a composition in which space and size are used to illustrate three-dimensional depth. This two-dimensional kind of representation intends to reflect the three-dimensional reality in which entities further away are perceived as relatively small and are located relatively high in the eye-range. However, this correspondence of two-dimensional depictions with the three-dimensional reality does not automatically result into correct figure-ground interpretations by all viewers. Several studies in the African context support the claim by Messaris (1994) and PATH (2002) that low skilled viewers have difficulties with pictorial conventions used for suggesting depth perspective. According to Bradley (1995), these studies, like the one conducted in the 1960’s by Holmes (referred to by Linney, 1995: 23) in Kenya and Hudson (referred to by Bradley, 1995: 74) in South Africa illustrate that graphic and environmental conventions need to be learned before ‘realistic’ visuals can be understood by unskilled viewers. Segall, Campbell & Herskovits (1963) found that people living on plains with a large horizon, who were familiar with the phenomenon that things become smaller as they are getting further way, understood the pictorial conventions concerning the depiction of figure-ground. People living in forests or high-rise cities on the other hand, were limited in their visual understanding of graphical conventions regarding figure-ground.
In sum, research has shown that static visuals, in particular the ones including conventional visual elements referring to abstract notions, may pose interpretation problems on different levels. Of central focus in the next section is the influence of education level in this context.

2.2.2 The influence of education level on the interpretation of static visuals

Literacy level and education level

In section 2.2.1, education level has been mentioned a few times as an influencing factor on the comprehension of visuals. In the studies where education level was factored in as a possible influencing factor (e.g. Carstens et al. 2006; Dowse & Ehlers, 2004; Foesenek, 2006), education level was used as an indicator for a respondent’s literacy level. Both concepts (literacy level and education level) and their relation will be discussed in this section. The aim of the section is to indicate the complexity of both concepts and to introduce the choices made with regard to the operationalization of education level in the current study (not to identify the exact influences of both concepts, on which the literature does not provide decisive insights).

As Carstens et al. (2006) and Maes, Foesenek & Hoogwegt (in press) point out, studying visuals in a low literate context is a challenging endeavour, as “the concept of low literacy is a capriciously amalgamated result of skills, experiences, knowledge, education and culture” (Maes et al., in press). As a consequence, a respondent’s literacy level is hard to measure. In the literature so far, education level (for which the years of formal schooling is commonly used as an indicator, e.g. by Carstens et al., 2006) is often used as an indicator for literacy level. Although it seems to make sense to assume that the lower (or higher) educated a respondent is, the lower (or higher) his or her literacy level is, one must be careful with ascribing results and effects to literacy level as such when only education level is factored in. Acquisition of literacy is only one of a wider range of (cognitive) skills acquired through formal schooling. This is shown by the nowadays considered classical study by Scribner & Cole (1981). Their findings run counter to the literacy hypothesis, an authoritative hypothesis until the early 80’s of the previous century (Kurvers, 2002) and followed by various authors (e.g. Finnegan, 1979; Goody, 1977; Goody & Watt, 1963; Havelock, 1963; Lenders, 1998; Levi-Strauss, 1962a; McLuhan, 1962, Ong, 1977). These authors suppose that the use of written language leads to changes in cognition and to the development of the modern, Western way of thinking (Kurvers, 2002). According to Kurvers (2002) these authors share the idea that the alphabetic script on its own has the ability to make changes in cognition. Vygotski (1962) added the idea that script does not only change the content of cognition (knowledge), but also a person’s manner of thinking (mental processes). For a more detailed overview of the ideas and approaches of these authors, see Kurvers (2002: 2-4). Studies by among others Luria (1976) and Greenfield (1972) were considered as empirical support for the literacy hypothesis. Scribner & Cole (1981) doubted to what extent the findings by Luria (1976) and Greenfield (1972) could be directly ascribed to literacy levels. According to them, the findings could equally well be the result of formal education, which was used to ‘measure’ the respondent’s literacy level. They conducted a replication study among the The Vai in Liberia. This population could serve as a ‘natural laboratory’, given that Vai, a syllabic script, was taught informally...
among adults. As a result, the Vai population included persons who acquired literacy without being exposed to formal education. Furthermore, two forms of schooled literacy exist in Liberia: Arabic, taught in Koran schools, as well as English, taught through formal education. Scribner & Cole (1981) compared five groups of adults: one group of illiterates and four groups of literates (in Vai, in Arabic, in Vai as well as Arabic and in English). Similar to Luria (1979) they used classification tasks as well as memory and reasoning tasks. No effects of literacy were found, while education did have an effect, e.g. for solving syllogisms (which was easier for the educated respondents). From their findings Scribner & Cole (1981) conclude that there is no evidence for the presupposed cognitive consequences of literacy: “our results are in indirect conflict with persistent claims that deep psychological differences divided literate and non-literate populations.” Although these studies do not focus on the interpretation of visuals, they are relevant for the present theoretical framework as they indicate that when only education level is taken into account, the effects found should not be automatically ascribed to literacy. Therefore, in the present study, where a person’s highest grade passed is used an indicator of education level, the terminology will only include education level and not literacy level (for example, respondents are be referred to as low educated instead of low literate). In this way the impression that effects can be ascribed to literacy level only, is avoided.

**Empirical support for the influence of education level on the interpretation of visuals**

The empirical studies outlined in section 2.2.1 reveal that (pictorial) conventions might pose interpretation problems for South African viewers. The results obtained in a number of these studies also reveal the effect of education level as an influencing factor on the interpretation of static visuals. Carstens *et al.* (2006) and Foesenek (2006) measured their respondent’s education level by taking into account their number of formal years of schooling. They found their low educated respondents to experience far more difficulties with the interpretation of (conventional) elements representing abstract notions (e.g. the mathematical symbols and thought balloons in figure 6), compared to the higher educated respondents. Furthermore, in all their studies among South Africans, Dowse & Ehlers (2001b, 2001a, 2003, 2004, 2005) found the comprehension of the pictograms to increase with formal years of schooling. Carstens *et al.* (2006) explain the relative low comprehension rates of their low educated respondents for the conventional symbols by arguing that these visual elements do not have any analogical residues which might trigger the meaning of relevant aspects of the visual. The differences between the low and high educated respondents were bigger for visual elements representing abstract notions without an analogical residue. Natural, analogical elements representing objects or humans did hardly pose any interpretation problems for both the low and high educated respondents in the studies by Carstens *et al.* (2006) and Foesenek (2006). Summarizing, Carstens *et al.* (2006) conclude that “a low level of formal education blocks the development of a vocabulary for abstract visuals” and Foesenek (2006) concludes that “abstraction has a negative effect on the comprehension of the intended message on visuals” and that this effect in particular applies to low educated viewers.

In the light of the studies referred to in section 2.2.1, the effect of education level is not surprising. After all, formal education results in exposure and familiarity with certain possible relevant
aspects resulting to the interpretation of visuals, e.g. familiarity with printed materials including visual conventions, as the studies referred to in this section indicate.

2.2.3 Summary

Summarizing, the studies discussed in section 2.2 show that pictures are not always worth a thousand words, i.e. they are not always worth the intended words. Furthermore, various studies have shown that respondents’ familiarity with visual elements (e.g. resulting from one’s own (bodily) experiences, culture or formal education) should be exploited in order to express the intended message. Especially for low educated audiences this should not be overlooked, as they are found to experience most difficulties in interpreting pictorial conventions invented to visually express abstract notions. For the context of health communication in particular, this is actually a sad finding, given that low educated audiences are also very much in need for information on HIV/AIDS and other health related topics. This is for instance shown by an overview of the quantitative research on the correlation between education level and HIV-infections in sub-Saharan Africa by the World Food Programme (2006). The overview leads to the conclusion that “higher levels of educational attainment are increasingly correlated with safer sexual behaviour and thus lower HIV prevalence rates” (p 13; see also Grosse & Auffrey, 1989; Kalichman et al. 1999; Weis & Coyne, 1997; Weis, Hart, McGee & D’Estelle, 1992 for the correlation between literacy level and health status). This implies that the audiences for which the comprehension of visuals in health communication materials is relatively low (i.e. the low educated audiences), are also the audiences for which the HIV prevalence rates are relatively high.

2.3 The arrow as an effective abstract visual element for suggesting motion in static visuals

The studies discussed in section 2.2 reveal that an effective representation of abstract notions in static visuals is quite a challenging endeavour. Familiar analogical, natural elements seem to be most successfully recognized and comprehended. However, sometimes the representation of abstract notions requires the use of abstract visual elements, simply because it seems impossible to effectively suggest the abstract notion by analogy. As mentioned before, the static nature of visuals is in sharp contrast with the dynamic, motion including actions, events or state of affairs one often intends to display by static visuals. Therefore, to suggest motion, an abstract visual element is often used, namely the arrow. This element can be seen as the prototypical symbol for the suggestion of motion. The effect and the interpretation of this element by viewers of static visuals is interesting, given (a) the frequently occurring intention for suggesting motion in static visuals, e.g. in health documents in (South) Africa and (b) the supposed capacity of the arrow to suggest motion and the direction of this motion in a visually efficient manner, as it is an uncomplicated visual element to display.
The arrow is a widely used symbol to indicate action. According to Krull & Sharp (2006) it is even the most frequently used indicator of action. Heiser & Tversky (2006) consider the arrow as a “compelling candidate for conveying change over time, movement and causality in a diagram”. Together with lines, boxes, crosses and circles, they belong to a class of diagrammatic elements (Tversky, 2005; Tversky et al. 2000). Arrows can represent asymmetric links between entities, indicating asymmetric relations, such as manner of movement, pointing or connecting, sequence, change over time, path or forces (e.g. Horn, 1998; Tversky et al., 2000). The arrowhead indicates the direction of the relation. In the context of suggesting motion in visuals this means that the arrowhead indicates the direction into which a certain movement takes place. Heiser & Tversky (2006) refer to analogical residues explaining the meaning and function of the arrow as an abstract element. They draw a parallel with the arrowhead used for hunting, leading the direction of the motion, and the V formed by water going downstream, showing the direction of motion of the water. According to Tversky (2002), the meaning of the arrow can possibly be inferred from this historical use or natural resemblance. She and her colleagues as well as other authors have conducted various experimental studies into the effect of the arrow in visuals. Contrary to most studies referred to in section 2.2, these studies took place in a Western context, with relatively high educated respondents. Results obtained in these empirical studies have confirmed the power of the arrow to indicate motion and the direction of this motion in static visuals.

Tversky, Zacks, Lee & Heiser (2000) studied the effect of arrows displayed in diagrams of complex systems on spontaneous interpretations of these illustrations. They found that when no arrows indicating motion and the direction of the motion are present in these illustrations, users interpret the content as just a group of static objects. However, when arrows are added to the illustration, users respond by expressions of a more dynamic nature: they try to understand how parts interact and they frame their understanding in terms of their personal role as agents of actions. Arrows were found to be interpreted as indicators of direction of motion. These findings are in line with several other earlier and later studies by Tversky and her colleagues. Heiser & Tversky (2006) examined the role of arrows in diagrams of mechanical systems, representing relatively complicated and unfamiliar scenes. During the first experiment, 67 participants described diagrams of mechanical systems (a car break, a bicycle pump and a pulley system) with or without arrows indicating the temporal sequence of the system. The obtained results show that the presence of arrows in the diagrams triggers spontaneous dynamic expressions: the participants primarily described the diagrams as conveying functional (dynamic) information. By the term functional, Heiser and Tversky (2006) refer to “the sequence of operational steps of the system”. What they consider as functional expressions are expressions of a dynamic nature, i.e. expressions including descriptions of motion of one or more of the elements displayed. Examples of such expressions are references to action or behaviour of the system (e.g. “The brake fluid travels down the tube”), as well as causal outcomes of the behaviour. The arrow was found to signal the direction, path and manner of movement. The diagrams without arrows on the other hand triggered spontaneous expressions including structural, static information (e.g. “The brake fluid is liquid”). The results obtained by this perception task indicate that the arrow triggers dynamic expressions. In addition to the perception task, Heiser & Tversky (2006) also
included a production task in their study. Based on either structural or functional descriptions, 196 participants drew sketches of mechanical systems. Arrows were spontaneously used to indicate functional processes in diagrams, making Heiser & Tversky (2006) conclude that “arrows can play a powerful role in augmenting structural diagrams to convey dynamic, causal, or functional information”. These findings about the interpretation and use of the arrow in perception and production tasks, are in line with earlier findings by Tversky & Lee (1999). They conducted an experiment in which the participants had to draw routes, making use of a toolkit containing elements which were found to be spontaneously used for this task in Tversky & Lee (1998). This toolkit included arrows (two types: straight and bent). The results show that 68% of the participants made use of arrows to indicate route directions. Earlier, Szichcinski (1979, 1980) had found similar results. He asked participants to draw visuals to explain how to operate a set of controls, such as toggle switches, rocker switches and rotary dials. It turned out that 96% of the participants used arrows to represent actions, and 44% used arrows combined with analogical elements: fingers or hands. Furthermore, Hegarty, Kriz & Cate (2003) conducted a study into people’s understanding of a complex device: a mechanical flushing cistern. They examined how well people can understand this machine by viewing a static diagram, by viewing a sequence of static diagrams showing different phases in the behaviour of the system (including arrows), by viewing computer animations and by viewing static or animated diagrams accompanied by verbal instructions. The motivation behind this study was the establishment that “the literature to date has failed to show any advantage of animations over static media for explaining how machines work” (Hegarty, Kriz & Cate, 2003). The research by among others Hearty (1992); Narayanan, Suwa & Motoda (1994, 1995); Schwartz & Black (1996) had already shown that people can be quite successful in inferring how a machine works from static diagrams, for instance when arrows indicate the movement of the system’s parts are present, as research by Heiser & Tversky (e.g. 2002, 2006) also revealed. This motion inferring process is called ‘mental animation’ by Hegarty et al. (2003). One of the findings by Hegarty et al. (2003) is that learning from animation is no better than learning from the three static diagrams showing different phases in the operation of the machine, accompanied by arrows.

Altogether, the studies show that, at least in the Western context, the abstract notion of motion – which contrasts with the static nature of visuals – can be effectively suggested by an abstract visual element, i.e. by the arrow.
2.4 Research questions and hypotheses

In this section the research questions and hypotheses following from the literature discussed in 2.1, 2.2 and 2.3 will be presented.

2.4.1 Main research questions and hypothesis

The literature referred to in section 2.3 indicates that the arrow has the capacity to suggest motion in visuals displaying relatively complex actions or sequences of motion. The question is to what extent these findings can be generalised to visuals aimed at low educated South Africans. Therefore, the first research question is:

RESEARCH QUESTION 1: Can motion in static visuals aimed at low educated South Africans be suggested by an abstract visual element (the arrow)?

As pointed out by Tverksy and colleagues, the arrow has the capacity to indicate the manner of motion, i.e. the specific direction of a motion. Therefore the present study will not only focus on the extent to which the arrow triggers dynamic expressions (henceforth referred to as D-expressions), as e.g. Tverksy & Heiser (2006) did. Of interest is also the extent to which the arrow has the capacity to indicate the direction of the motion and as a result triggers the correct dynamic expression (henceforth referred to as CD-expression). Therefore a second research question, relating to the direction of motion, is formulated:

RESEARCH QUESTION 2: Can the direction of motion in static visuals aimed at low educated South Africans be suggested by an abstract visual element (the arrow)?

It is not completely self-evident that the findings by Tversky and others can be generalized to a low educated population in a non-Western context. The South African population differs from the ones in the studies referred to above, for instance in terms of education level, culture, (social economic) environment, etc. The arrow as a symbol is not a natural, analogical visual element. On the other hand, it is not an arbitrary, purely conventional symbol either: it has analogical roots in reality from which its meaning can possibly be inferred according to Heiser & Tversky (2006). Given the arrow’s capacity to indicate motion and direction in illustrations of a relatively complex nature, the following hypothesis can be formulated:

6 The CD-expressions form a subset of the D-expressions. A D-expression is labelled as a CD-expression when in this expression it is referred to the motion indicated by the arrow (also referred to as ‘the intended motion’). To avoid confusion: this does not mean that non-correct (D-)expressions are ‘wrong’ and/or do not make any sense. They just do not include a reference to the motion which is indicated by the arrow.
HYPOTHESIS 1: Arrows in static visuals are effective indicators of motion and the direction of motion for low educated South Africans, i.e. they trigger spontaneous D- and CD-expressions.

2.4.2 Additional research question and hypotheses

The beneficial effect of arrows is in particular expected for visuals displaying relatively unfamiliar actions. For these visuals it can be expected that the viewer has to rely to a great extent on the arrow to interpret the visual in terms of (a specific) motion, given the unfamiliarity or lack of knowledge by the viewer with regard to what is displayed, resulting in no or hardly any other displayed visual cues apart from the arrow (if present) for motion on the direction of motion.

HYPOTHESIS 2: The beneficial effect of arrows occurs to a greater extent in visuals displaying relatively unfamiliar actions than for visuals displaying relatively familiar actions.

The hypothesis with regard to the influence of education level following from the literature overview (e.g. the studies by Carstens et al., 2006; Dowse & Ehlers, 2004; Foesenek, 2006) is the following:

HYPOTHESIS 3: Arrows trigger more D and CD-expressions as education level increases.

This hypothesis is based on the assumption (supported by the research referred to in section 2.1.2.2) that the higher a respondent is educated, the more likely he or she is to be familiar with (partly) conventional visual elements, including the arrow. Again, the effect is in particular expected to be the case for visuals displaying relatively unfamiliar actions in which few other cues for (the intended) motion apart from the arrow (if present).

Age is taken into account as well, given that the older one is, the longer ago it probably has been that the formal education took place, leading to the following hypotheses:

HYPOTHESIS 4: Arrows trigger more D and CD-expressions as age decreases.

Again, for the same reasons as formulated for the previous hypotheses the influence of age on the type of expressions by the respondents is expected to be revealed most clearly by visuals displaying relatively unfamiliar actions.

It is interesting to know to what extent the CD-expressions are based on the power of the arrow to indicate a specific motion, or perhaps on others elements in the visual. This leads to the following, explorative, research question:

RESEARCH QUESTION 3: Are there any other visual elements apart from the arrow that have the capacity to reveal (the direction of) motion?
Throughout all the results, attention will be paid to the influence of a second response moment (elicitation), providing a second opportunity for the respondent to come up with a response, in which he or she is guided, or directed, toward D- and CD-expressions. Given its guiding nature, this elicitation is expected to bring the respondents exposed to one of the visual variant (visuals including arrows versus visuals without arrows) and the restively lower and higher educated respondents closer together in terms of (C)D-expressions. Of course, the spontaneous expressions are of greatest interest, given that they are the most natural responses. Moreover, from a practical point of view the spontaneous responses are most interesting as well, as in reality there will probably not be a second person who will guide the viewer towards a more extensive interpretation than the initial one.

2.5 Method

In the current section the method of Study 1 will be discussed. First of all, relevant details about the respondents will be presented in section 2.5.1. Following, the research materials will be presented in section 2.5.2. In section 2.5.3, the setting of the interviews will be explained, followed by a description of the procedure of the interview sessions in section 2.5.4. The design is presented in section 2.5.5 and the analyses will be outlined in the last section, 2.5.6.

2.5.1 Respondents

Fifty respondents, all speaking Sepedi or Setswana, were interviewed. The interviews with two respondents could not be analyzed because of a technical error during the recording of the interviews, resulting in a recording with no audio information. Two other respondents indicated to be poor-sighted. They were left out of the analyses as well, as they might have been unable to see the research materials properly. The features of the remaining 46 respondents can be found in table 2.1.

<table>
<thead>
<tr>
<th>Visual variant</th>
<th>n</th>
<th>Grade</th>
<th>Age</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>- arrow</td>
<td>23</td>
<td>7.26</td>
<td>32.48</td>
<td>47.8%</td>
<td>52.2%</td>
</tr>
<tr>
<td>+ arrow</td>
<td>23</td>
<td>7.48</td>
<td>30.65</td>
<td>43.5%</td>
<td>56.5%</td>
</tr>
<tr>
<td>Overall</td>
<td>46</td>
<td>7.37</td>
<td>31.57</td>
<td>45.7%</td>
<td>54.3%</td>
</tr>
</tbody>
</table>

NB: ‘-arrow’ refers to the control condition including the visual variants without any arrows, while ‘+arrow’ refers to the experimental condition including the visual variants with arrows.

The respondents were randomly assigned to one of the visual variants (+arrow versus -arrow). Education level (measured in grades\(^7\)) and age are based on self-reports of the respondents. It should

\(^7\) If no classes or doubled, a person’s grade equals the years of schooling this person has had.
be taken into account that – presumably to save face – the grade reported by low literates is often adjusted upward by a few levels (Carstens et al., 2006). Furthermore, a few respondents were not able to recall their exact grade and/or age, and they therefore had to give an approximation. The respondents were all inhabitants of urban townships in Pretoria. The majority of the respondents lived in the township Atteridgeville. Each respondent received a reward of twenty Rand (approximately three euros) for his or her participation.

2.5.2 Research materials

Preselection of the research materials

The initial research materials comprised eighteen visuals of which most were selected from Soul City’s health communication brochures (see appendix A). The materials were selected from these brochures in order to stay as close as possible to the materials which are actually aimed at the research population. Moreover, in this way the findings of the present study may yield some recommendations about real existing South African communication materials. The selected visuals are all designed to show an uncomplicated – but not necessarily familiar – action performed by a human (the actor). The presence of the actor is suggested by depicted hands and/or the human body. The actions intended to be displayed are of a dynamic nature, i.e. they imply the spatial movement of one or more elements in the visual. The selection for instance includes a depiction of a lid put on a pot by human hands and a male kicking a soccer ball. Two variants of each visual were created. The experimental condition comprises the eighteen visuals to which an arrow is added. This arrow indicates a specific spatial motion (also referred to as ‘the intended motion’) of (an) element(s) in the visual. The control condition includes the same visuals, with no arrows present.8 Pilot interviews were held with nine respondents in The Netherlands. The details about these pilot interviews can be found in appendix C.

Final selection of the research materials

The final selection of visuals comprises eleven visuals (see appendix B). The visuals are expected to vary in familiarity. Depictions including familiar elements relating to familiar actions (henceforth referred to as ‘Easy Visuals’) can be expected to evoke (C)D-expressions to a relatively great extent. These visuals are less likely to reveal any effects of the arrow, and therefore visuals displaying less familiar actions (henceforth referred to as ‘Difficult Visuals’) were included as well, as they are expected to evoke (C)D-expressions to a smaller extent.

Easy Visuals First of all, the visuals Fighting Couple, Condom and Package, Soap and Bowl and Cigarettes and Dustbin are expected to suggest motion (i.e. to trigger D-expressions) to a relatively great extent. These visuals relate to actions in the (daily) domestic environment and topics which are

8 The original versions of the visuals did not include any arrows. However, the visual Asthma Pipe is an exception. The original version of this visual is the visual in the +arrow condition. The upper arrow represents pressure of the finger, implying an actual movement of the finger (and the upper part of the pipe), the lower arrow only represents pressure and no motion. Given the interest of the present study in the arrow as an indicator of (a specific) motion, only the upper arrow was taken into account in the analyses.
prototypical for South African health communication materials. Given that these topics are usually referred to in a non-static, educational and advisory way, they can be expected to be familiar, and semantically relatively rich in the eyes of the viewers. The actions suggested by the visual Pot and Lid as well as Salt and Bowl are expected to be familiar as well, given that the depicted elements and the action they suggest are prototypical for activity of cooking, an activity daily performed all over the world. They are expected to trigger familiar bodily experiences. Moreover, the law of gravity can be expected to reveal the downward movement of the salt in the visuals Salt and Bowl (which is of course also the case for the visual Cigarettes and Dustbin).

**Difficult Visuals** Though also relating to the domestic environment, the visuals Asthma Pipe and Bottle and Pipe are expected to evoke motion to a smaller extent. They are expected to be less familiar, given that they intend to suggest actions which are not necessarily generally known. It seems for instance unlikely that many viewers will be familiar with the existence or working of an asthma pipe, given that this medical instrument is simply not a part of the daily life of many persons. Furthermore, the objects in the visual Bottle and Pipe (which is intended to show how an home-made asthma pipe can be constructed) are a not very frequently occurring combination and are therefore not expected to spontaneously remind many of the viewers of the intended activity. Finally, the original versions of the visuals depicting gymnastic exercises are also expected to trigger D-expressions to a relatively small extent, as no cues for (the intended) motion are present. As a result, if the arrow would have any effects, the expressions triggered by these visuals and the visuals Asthma Pipe and Pipe and Bottle can be expected to show this most clearly.

**Classification of the visuals for further analyses**

The distinction between Easy and Difficult Visuals should be reflected in the number of spontaneous (C)D-expressions. The table below displays the proportion spontaneous D-expressions, presented per visual, for the –arrow visual variant.

<table>
<thead>
<tr>
<th>Visual</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soap and bowl</td>
<td>1.00</td>
</tr>
<tr>
<td>Cigarettes and dustbin</td>
<td>.95</td>
</tr>
<tr>
<td>Fighting couple</td>
<td>.91</td>
</tr>
<tr>
<td>Condom and package</td>
<td>.87</td>
</tr>
<tr>
<td>Salt and bowl</td>
<td>.87</td>
</tr>
<tr>
<td>Pot and lid</td>
<td>.78</td>
</tr>
<tr>
<td>Bottle and pipe</td>
<td>.61</td>
</tr>
<tr>
<td>Asthma pipe</td>
<td>.39</td>
</tr>
<tr>
<td>Exercise – sit up straight</td>
<td>.30</td>
</tr>
<tr>
<td>Exercise – sit and spread arms</td>
<td>.26</td>
</tr>
<tr>
<td>Exercise – standing</td>
<td>.09</td>
</tr>
</tbody>
</table>

The table shows that the variation among the visuals is substantial. Furthermore, it reveals that the sequence of the visuals is in line with the expectations as outlined in the previous section. For the visuals that are highly likely to evoke D-expressions even without the arrow, a ceiling effect may occur as hardly any opportunity is left for the arrow to cause any effects. These visuals are highly unlikely to reveal effects of the arrow. On the other hand, the visuals evoking dynamic expressions to a relatively small extent are likely to reveal possible effects. Therefore, in further analysis, the visuals will be clustered in these two groups. Two visual types with corresponding levels of difficulty are distinguished: visuals evoking D-expressions to a
relatively strong extent (Easy Visuals) and visuals evoking D-expressions to a relatively weak extent (Difficult Visuals). The dividing line between both clusters used for classifying the visuals is specified according to the following procedure. A visual is considered to be Easy when it is ‘sure’ that in half of the cases the visual triggers a D-expression. In other words: the score for a D-expression lies above the 95% confidence interval of random expressions (p=.50). The upper limit of the confidence interval is .70\(^9\), so all visuals yielding proportions above .70 are considered to be Easy. This leads to a dividing line between the visual Pot and Lid (.78) and Bottle and Pipe (.61). The proportions D-expressions that are correct, show the exact same sequence of visuals from easy to difficult as the proportions D-expressions.

**2.5.3 Interview setting**

Most interviews were held in a quiet room in a health clinic in the township Atteridgeville. During each interview session, an interpreter and the researcher were present. All interviews were recorded with a small microphone and a webcam. A more detailed description of the interview setting can be found in appendix D.

**2.5.4 Procedure**

Each interview session started with an introduction, in which the respondent was explained that the goal of the interview was to see whether the visuals he or she was going to be shown, would be clear enough to use in public information documents.

In order to create a context for the respondent, the visuals were preceded by a short introduction. The first group of four visuals (including respectively Pot and Lid, Salt and Bowl, Bottle and Pipe and Asthma Pipe) were introduced as images relating to the domestic environment and were presented to the respondent one by one. The procedure per visual was as follows. In order to see to what kind of spontaneous expressions the visual triggers, the respondent was first asked a neutral question: “What do you see on this visual?”. In case the expression after this first response moment was not of a dynamic nature, i.e. was no D-expression, a second response moment was created to gain insight into whether more guidance (elicitation) would as yet lead to a (possibly correct) D-expression. For this purpose, the respondent was asked whether he or she saw any motion in the visual. In the case of a CD-expression (either spontaneously or elicited) the respondent was asked the follow-up question: “What makes you think that?”. The second group of visuals (including the three Exercise visuals) was introduced as gymnastic exercises and the procedure per visual was exactly similar to the one for the first group of visuals. The last cluster included the visuals Fighting Couple, Condom and Package, Soap and Bowl and Cigarettes and Dustbin. These visuals were introduced as images about the domestic environment from which the respondent might learn something. The

\[^9\] The standard deviation for n=23 is \(\sqrt{(p^*q)/n}\), \(\sqrt{(0.50 * 0.50 / 23)} = .104\). The upper limit of the confidence interval is .50 + 1.96*.104 = .70
question “What can you learn from this picture?” was added to the procedure per visual in order to see whether these visuals were indeed of a normative and educating nature. In case no explicit comments on the arrow had been made spontaneously by the respondents exposed to the +arrow visual variant, he or she was explicitly asked after its meaning at the end of the interview. He or she was once again shown the visual Exercise – sit up straight. The interviewer pointed to the arrow and asked “Do you know what this is?” This procedure was applied with one of the Exercise-visuals in order to make sure that (a) the respondent was exposed to the arrow in the context of motion indication in visuals and that (b) it concerned a visual of which chances are as small as possible that (the intended) motion – and with that possibly the function of the arrow – can be inferred from visual elements other than the arrow.

More details on the procedure, and the insights yielded by the pilot studies, can be found in appendix E.

2.5.5 Design

The effect of Visual Variant (+arrow versus –arrow) is measured by means of a between subjects design. Furthermore, the experiment uses a repeated measurements design with Response Moment (with levels: spontaneous; spontaneous + elicited) as a within-subject factor, and the proportions D-expressions and CD-expressions (treated as a subset of the D-expressions) as the dependent variables.

2.5.6 Analyses

Transcript and Kappa-scores

The verbal expressions of each respondent were literally transcribed and classified per visual variant (+ arrow versus -arrow), visual and question. Relevant visual information was included in the transcript as well. It was for instance indicated when a respondent gestured the intended motion while he or she did not mention it verbally. The transcription served as a basis for the coding of the data. The data were coded by hand by a first analyst. A second analyst coded 40% of the answers with regard to the D- and CD-expressions. This second analysis was conducted because analyzing the responses with regard to these focal points might involve some subjectivity. The resulting Kappa scores of the two analysts are the following:

- Static versus Dynamic expression (D-expression): .87
- For D-expression: CD-expression versus no CD-expression: .95

10 The answers to this question by the respondents showed that this was indeed the case, but did not yield any interesting and relevant insights in the light of the research questions. Therefore, these results will not be included in the reports of the results.

11 This occurred only sporadically and if so, it mostly concerned the visual Asthma Pipe.
The majority of differences between the two analyses appeared to be caused by slight structural differences in the applied definitions of both concepts by the analysts. A short discussion between both analysts resulted in a final overall agreement.

Coding of the transcript

In this section the basic coding scheme will be briefly outlined. More detailed information on the coding of the transcript can be found in appendix F.

1. The spontaneous and elicited responses to the visuals, were analyzed on the expression type:
   - Static expression versus D-expression;
   - For D-expressions: CD-expression versus no CD-expression

2. In case the respondent came up with a CD-expression, the source he or she identified for this expression was analyzed. After a qualitative analysis of the responses, the whole range of expressions was divided into the following categories and subcategories:
   - Abstract visual element: arrow
   - Analogical visual element
     - Bodily element
       - Hand(s)
       - Body
     - Object

3. For each respondent exposed the +arrow visual variant, it was analyzed how he or she interpreted the arrow (spontaneously during the interview, or elicited, after the explicit question about the meaning of the arrow). A qualitative analysis resulted in the following categories:
   - Indicator of motion, and the direction of this motion in visual (the intended interpretation)
   - Indicator of direction (not further specified)
   - Respondent is unfamiliar with arrow

Statistical analyses

All statistical analyses were performed with the help of SPSS. To test for significant effects regarding the Difficult and Easy visuals, the following analyses were performed for both levels of difficulty separately.

To test for significant effects of Visual Variant (+ arrow versus - arrow) and Response Moment (spontaneous versus spontaneous + elicited) on the proportions D-expression, a two-way analysis of variance for repeated measures (one-tailed) was performed with Visual (either the five Difficult or the six Easy Visuals) and Response Moment (spontaneous versus spontaneous + elicited) as the within subject-factors and Visual Variant as a between subject factor. Given that the CD-expressions form a subset of the D-expressions, each proportion CD-expressions is of a conditional nature. It concerns the proportion CD-expressions among the D-expressions. Therefore, each respondent’s proportion CD-expressions, is calculated over visuals triggering D-expressions for the specific respondent only. Given that the number of visuals which should be taken into account (i.e. the number of visuals triggering D-expressions) varies among the respondents, for each respondent one overall proportion
CD-expressions was calculated for the Easy Visuals, and one for the Difficult Visuals. To test for significant differences regarding the proportions D-expressions including a CD-expression, a one-way-analysis of variance for repeated measures (one-tailed) was performed with Response Moment as a within subject-factor and Visual Variant as the between subject factor.

To test for significant effects of Visual Variant \textit{within a response moment}, a univariate analysis was applied to these proportions, with Visual Variant as the independent variable and the proportion D-expressions and the proportion CD-expressions as the dependent variables.

Finally, Pearson correlation analyses (one-tailed) were conducted to identify whether there was any relationship between Education Level and the proportion D-expressions, as well as between Education Level and the proportion CD-expressions. The same analysis was applied to test for the effect of Age\textsuperscript{12}.

2.6 Results

In section 2.6.1, the results regarding the effect of the arrow on the interpretations of the visuals are presented. Subsequently, an overview of the various interpretations of the arrow by the respondents is presented in section 2.6.2. Thirdly, in section 2.6.3 the focus will be on the explanatory power of Education Level and Age on these results. At last, the results regarding the sources for CD-expressions identified by the respondents are presented in section 2.6.4.

Given that the spontaneous responses\textsuperscript{13} are of greatest interest, as well as the responses yielded by Difficult Visuals\textsuperscript{14}, the proportions relating to these visuals and responses are highlighted in the tables. To illustrate this: when the proportion for a CD-answers is .18, this means that 18\% \textit{of the D-expressions} was correct.

2.6.1 Effect of arrow on expression type

First of all, the results regarding effect of the arrow on the proportion D-expression will be presented. Subsequently, the results with respect to the proportions CD-expression will be of central focus in this section. It is important to take into account that the proportion correct dynamic (CD) expressions is a subset of the proportion dynamic (D) expressions. Each proportion CD-expressions relates to the amount of D-expressions including the correct dynamic expression, i.e. the expression including the intended motion indicated by the arrow.

\textsuperscript{12} Neither for the -arrow or the +arrow condition, correlations between Age and Education Level were found (-arrow: \(r=-.32, p=.14\); +arrow: \(r=-.14; p=.46\)) with a Pearson correlation analysis.

\textsuperscript{13} Because of their pure, unguided nature

\textsuperscript{14} Given their higher probability capacity to reveal effects of the arrow than the Easy Visuals
Dynamic expressions

Table 2.4 presents the proportion D-expressions in relation with Level of Difficulty (difficult versus easy), Response Moment (spontaneous versus spontaneous + elicited) and Visual Variant (+arrow versus -arrow). The proportions under ‘Spontaneous’ relate to the spontaneous expressions given in the first response moment. The proportions under ‘Spontaneous + Elicited’ include the spontaneous expressions as well as the expressions after elicitation. These proportions comprise both response moments, and show to what extent the elicitation (the second response moment) resulted in an increase of the proportions. This is also the case for all following tables including the factor Response Moment.

Table 2.4 Proportion D-expressions in relation with Level of Difficulty, Response Moment and Visual Variant

<table>
<thead>
<tr>
<th></th>
<th>Spontaneous</th>
<th>Spontaneous + Elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- arrow (n=23)</td>
<td>+ arrow (n=23)</td>
</tr>
<tr>
<td>Difficult</td>
<td>.33</td>
<td>.50 ***</td>
</tr>
<tr>
<td>Easy</td>
<td>.90</td>
<td>.86 ns</td>
</tr>
</tbody>
</table>

|           | - arrow (n=23)    | + arrow (n=23)          |
| Difficult | .71               | .75 ns                  |
| Easy      | .96               | .96 ns                  |

NB: * : p<.05,  ** : p<.025,  *** : p<.01,  ****: p<.001,  ns : not significant

For Difficult Visuals, overall effects were found for Visual Variant ($F(1,44)=7.41$, $p<.01$, $\eta^2=.14$) and Response Moment ($F(1,44)=113.37$, $p<.001$, $\eta^2=.72$). D-expressions occurred more often for visuals with an arrow (.63 versus .52), and for the spontaneous and elicited responses combined (.73 versus .42). There was an interaction between Visual Variant and Response Moment ($F(1,44)=5.60$, $p<.025$, $\eta^2=.11$). Split analyses showed that the effect of Visual Variant had to be attributed to the spontaneous expressions ($F(1,44)=9.82$, $p<.025$, $\eta^2=.18$; spontaneous + elicited responses: $F(1,44)=.76$, $p=.19$). In spontaneous reactions, the absence of an arrow lowered the proportion D-expressions (.33 versus .50). When the elicited responses are taken into account as well however, the no-arrow visuals had caught up almost completely (.71 versus .75).

For Easy Visuals, an overall effect was found for Response Moment ($F(1,44)=20.28$, $p<.001$, $\eta^2=.32$; Visual Variant: $F(1,44)=.38$, $p=.23$; RM*DV: $F(1,44)=.96$, $p=.33$). In spontaneous responses the proportion D-expressions was slightly lower (.88 versus .96). Whether an arrow was present or not, had not influenced the type of response.

Correct Dynamic expressions

Table 2.5 presents the proportion CD-expressions in relation with Level of Difficulty, Response Moment and Visual Variant.
Table 2.5 Proportion CD-expressions in relation with Level of Difficulty, Response Moment and Visual Variant

<table>
<thead>
<tr>
<th></th>
<th>Spontaneous</th>
<th>Spontaneous + Elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- arrow</td>
<td>+ arrow</td>
</tr>
<tr>
<td></td>
<td>(n=23)</td>
<td>(n=23)</td>
</tr>
<tr>
<td>Difficult (n=5)</td>
<td>.18</td>
<td>.31 *</td>
</tr>
<tr>
<td>Easy (n=6)</td>
<td>.81</td>
<td>.79 ns</td>
</tr>
<tr>
<td></td>
<td>.42</td>
<td>.54 *</td>
</tr>
<tr>
<td></td>
<td>.91</td>
<td>.90 ns</td>
</tr>
</tbody>
</table>

NB: * : p<.05,  ** : p<.025,  *** : p<.01,  ****: p<.001,   ns : not significant

For Difficult Visuals, overall effects were found for Visual Variant ($F(1,44)=3.84, p<.05, \eta^2=.08$) and Response Moment ($F(1,44)=56.26, p<.001, \eta^2=.56$). D-expressions were correct more often for visuals with an arrow (.43 versus .30) and for the spontaneous and elicited responses combined (.25 versus .48). There was no interaction between Visual Variant and Response Moment ($F(1,44)=.023; p=.44$).

For Easy Visuals, an overall effect of Response Moment was found ($F(1,44)=35.99, p<.001, \eta^2=.45$); Visual Variant ($F(1,44)=.12; p=.37$; RM*DV: $F(1,44)=.01, p=.45$). In spontaneous responses the proportion CD-expressions among the D-expressions was slightly lower (.80 versus .91). It appears that the presence of an arrow had not influenced the type of response.

2.6.2 Interpretation of arrow

For each respondent exposed to the +arrow visual variant, it was examined how he or she interpreted the arrow, spontaneously (during the interview) or elicited (after the explicit question). All respondents who were able to and had the chance to come up with an interpretation, interpreted the arrow as an abstract element. Furthermore, approximately half of the respondents is familiar with the meaning of the arrow as indicator of motion in visuals (52.2%; 43.5% mentioned the meaning spontaneously and 8.7% did this after elicitation), while 17.4% of the respondents is not familiar with the arrow at all. A percentage of 13.0% did not mention the arrow spontaneously and – after elicitation – identified it as an indicator of direction, without any further specifications. By mistake of the researcher, 17.4% of the respondents was not asked the question after the meaning of the arrow, even though they had not mentioned it spontaneously during the interview.

2.6.3 Effect of education level and age on expression type

In the current section, the results regarding the influence of the factors Education Level and Age will be presented. First of all, the influence of education level and age on the proportion a D-expressions will be of central focus. Subsequently, the correlations regarding the proportion D-expressions including the CD-expression will be presented.
Influence of Education Level and Age on proportion D-expressions

The correlations presented in Table 2.6 show to what extent education level and age correlate with the proportion D-expressions (presented in relation to Level of Difficulty, Response Moment and Visual Variant).

Table 2.6: correlations between Education Level and proportion D-expressions, as well as Age and proportion of D-expressions, in relation to Level of Difficulty, Visual Variant and Response moment

<table>
<thead>
<tr>
<th></th>
<th>Spontaneous</th>
<th>Spontaneous + Elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- arrow</td>
<td>+ arrow</td>
</tr>
<tr>
<td></td>
<td>((n=23))</td>
<td>((n=23))</td>
</tr>
<tr>
<td>Ed. level</td>
<td>Difficult ((n=5))</td>
<td>.11 ns</td>
</tr>
<tr>
<td></td>
<td>Easy ((n=6))</td>
<td>-.09 ns</td>
</tr>
<tr>
<td>Age</td>
<td>Difficult ((n=5))</td>
<td>.06 ns</td>
</tr>
<tr>
<td></td>
<td>Easy ((n=6))</td>
<td>.21 ns</td>
</tr>
</tbody>
</table>

NB: * : p<.05, ** : p<.025, *** : p<.01, ****: p<.001, ns : not significant

The table shows that spontaneous expressions as a response to Difficult Visuals including arrows, correlate with Education Level, i.e. they can be (partly) explained by this factor. The higher a respondent’s education level, the higher the proportion spontaneous D-expressions for these visuals. This is not the case for the other scores of visual variant with arrows. Instead, they correlate negatively with Age, indicating that the younger a respondent, the higher the proportion D-expressions. (The non-significant correlations with regard to Education Level and Age: respectively p=.47 to .12 and p=.39 to .11).

Effect of education level and age on proportions D-expressions and CD-expressions

The correlations presented in Table 2.7 show to what extent Education Level and Age correlate with the proportion correct dynamic expressions (presented in relation to Level of Difficulty, Response Moment and Visual Variant).

Table 2.7: correlations between Education Level and proportion CD-expressions, as well as Age and proportion CD-expressions, in relation to Level of Difficulty, Visual Variant and Response Moment

<table>
<thead>
<tr>
<th></th>
<th>Spontaneous</th>
<th>Spontaneous + Elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- arrow</td>
<td>+ arrow</td>
</tr>
<tr>
<td></td>
<td>((n=23))</td>
<td>((n=23))</td>
</tr>
<tr>
<td>Ed. level</td>
<td>Difficult ((n=5))</td>
<td>.20 ns</td>
</tr>
<tr>
<td></td>
<td>Easy ((n=6))</td>
<td>.38 *</td>
</tr>
<tr>
<td>Age</td>
<td>Difficult ((n=5))</td>
<td>-.12 ns</td>
</tr>
<tr>
<td></td>
<td>Easy ((n=6))</td>
<td>-.50 ***</td>
</tr>
</tbody>
</table>

NB: * : p<.05, ** : p<.025, *** : p<.01, ****: p<.001, ns : not significant
When it comes to Easy Visuals without any arrows, both Education Level and Age were found to correlate with the proportion spontaneous CD-expressions. The higher educated and the younger a respondent is, the more often his or her D-expression is the correct one. When also taking into account the elicited responses, education level and age do no longer exert any influence on the expression type. (The non significant correlations with regard to Education Level and Age: respectively p=.45 to .15 and p=.18 to .38).

2.6.4 Sources for correct dynamic expressions

In case a respondent came up with the CD-expression, he or she was asked what made him or her think that the specific motion was displayed. The various categories of the resulting sources for the CD-expression identified by the respondents can be found in the two tables below (Table 2.8 and 2.9). These tables present the proportions for each source. The results are presented for two groups of visuals separately. The visuals are categorized by the visual elements they include (i.e. by the visual elements that could possibly be identified as the source for the CD-expression).

1. Visuals with full body (i.e. not only hands) and without objects (table 2.9):
   - Exercise – Sit up Straight
   - Exercise – Sit and Spread Arms
   - Exercise - Standing
   - Fighting Couple

2. Visuals with objects and hands (table 2.10):
   - Pot and Lid
   - Salt and Bowl
   - Bottle and Pipe
   - Asthma Pipe
   - Condom and Package
   - Soap and Bowl
   - Cigarettes and Dustbin

For each respondent the mean proportion for each source was calculated, from which an overall mean proportion per source for all respondents could be calculated. Given that it concerns a source for the CD-expression, the proportion per respondent is based on the visuals triggering a CD-expression only. As a result, the number of visuals on which the proportion for a certain source is based, may vary among the respondents. For several respondents the proportion could not be calculated. For none of the visuals to which these respondents were exposed, sources for the CD-expression are known.
Either because he or she did not come up with a CD-expression, or because of a missing value\textsuperscript{15} with regard to the source for the CD-expression.

In order to test for possible correlations with Education Level, Pearson correlation analyses were applied to the proportions. Only significant relations are reported.

Table 2.8: Proportions identified sources for CD-expression for the visuals with full body, in relation to Visual Variant

<table>
<thead>
<tr>
<th>Source Type</th>
<th>+Arrow (n=20)</th>
<th>-Arrow (n=11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract: Arrow</td>
<td>.26</td>
<td>NA</td>
</tr>
<tr>
<td>Abstract + analogical: Arrow + Body</td>
<td>.02</td>
<td>NA</td>
</tr>
<tr>
<td>Analogical: Arrow + Body</td>
<td>.68</td>
<td>.82</td>
</tr>
<tr>
<td>Hand</td>
<td>.25</td>
<td>.55</td>
</tr>
<tr>
<td>Hand + Body</td>
<td>.43</td>
<td>.27</td>
</tr>
<tr>
<td>No source identified</td>
<td>.05</td>
<td>.18</td>
</tr>
</tbody>
</table>

NB: NA = not applicable

The table shows that the proportion for arrows is .26. This proportion was found to correlate with Education Level ($r=.40$, $p<.05$), revealing that the higher a respondent's Education Level, the higher the proportion responses in which the arrow was identified as a source for the CD-expression. With a proportion of .68, analogical elements are more likely to be identified as the source for the visuals including arrows than the arrow. What cannot be read from the table but is interesting to report is that when a respondent responded with a CD-expression to one of the visuals depicting gymnastic exercises, in all cases the arrow was identified as the source for this expression. This is in line with the earlier formulated expectation that for these visuals the arrow is the only possible source that might lead the viewer to a CD-expression. However, given that only five respondents came up with a CD-expression for one or more of these visuals, this proportion is calculated over these five respondents only. Table 2.9 presents the results for the visuals with objects and without the human body.

\textsuperscript{15} For about a quarter of the ID-expressions (23%), the respondent was not asked after his/her motivation, resulting in missing values. The proportions in table 2.8 and 2.9 are based on the remaining 77% of the ID-expressions. Relatively many missing values in the –arrow condition occur for the visual Fighting Couple (55%). Added to the fact that the other human body including visuals (the exercises) evoked only one ID-expression in the –arrow condition, relatively few respondents ($n=11$) are taken into account in the –arrow condition for the visuals including the human body.
Table 2.9: proportions of identified sources for CD-expression for the visuals including objects and hands, in relation to Visual Variant

<table>
<thead>
<tr>
<th>Source</th>
<th>+ARROW (n=22)</th>
<th>-ARROW (n=22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract: Arrow</td>
<td>.21</td>
<td>NA</td>
</tr>
<tr>
<td>Abstract + analogical</td>
<td>.08</td>
<td>NA</td>
</tr>
<tr>
<td>Arrow + Hand(s)</td>
<td>.06</td>
<td>NA</td>
</tr>
<tr>
<td>Arrow + Object(s)</td>
<td>.02</td>
<td>NA</td>
</tr>
<tr>
<td>Analogical</td>
<td>.56</td>
<td>.72</td>
</tr>
<tr>
<td>Object(s)</td>
<td>.17</td>
<td>.20</td>
</tr>
<tr>
<td>Hand(s)</td>
<td>.31</td>
<td>.46</td>
</tr>
<tr>
<td>Object(s)+Hand(s)</td>
<td>.07</td>
<td>.06</td>
</tr>
<tr>
<td>No source identified</td>
<td>.15</td>
<td>.27</td>
</tr>
</tbody>
</table>

NB: NA = not applicable

The results in the table reveal that the proportions of responses in which an arrow was identified as the source for the CD-expression is .21. The proportion responses in which an arrow was identified as the source together with an analogical element (hand or object) is .08. Both combinations were found to correlate with grade (arrow and object: r=.48, p < .025; arrow and hands: r=.40, p<.05). With a proportion of .56 the analogical elements are again more likely to be identified as the source for the CD-expression than the arrow. Hands are most likely to be identified as the source. This is also the case for the visuals without arrows.

2.7 Conclusions and Discussion

In this section conclusions will be drawn from the results presented in section 2.6, and they will be briefly discussed (a more extensive discussion, also including the findings obtained in Study 2, is outlined in chapter 4). For the reasons outlined before, the main focus will be on the spontaneous responses.

2.7.1 Effect and interpretation of the arrow

Is the arrow an effective indicator of motion and the direction of motion in static visual aimed at low educated South Africans? The results show it is. When present in Difficult Visuals, the arrow results in more spontaneous dynamic expressions compared to when no arrow is present (even though the visuals displaying gymnastic exercises were introduced as such, implying motion including activities being performed by the human on the visual). Moreover, these spontaneous dynamic expressions are correct more often (i.e. include more often the intended motion indicated by the arrow) when an arrow is displayed in the visual. This effect even remains when the elicited dynamic expressions are taken into account as well. These results indicate that the arrow not only has the expressive power to
suggest motion, but also the direction of this motion. These effects being found for Difficult Visuals, demonstrates that the beneficial effect of the arrow indeed in particular applies to visuals displaying relatively unfamiliar actions. As expected, the depictions on the Easy Visuals were so familiar to the respondents, that they were 'too easy' to reveal any effects of the arrow. In accordance with the expectations, the elicitation indeed guided the respondents towards an expression in terms of (the intended) motion.

The results are consistent with previous research (e.g. Hegarty, 1992; Hegarty, Kriz & Cate, 2003; Heiser & Tversky, 2006; Narayanan, Suwa & Motoda, 1994, 1995; Schwartz & Black, 1996; Szichcinski, 1979, 1980; Tversky & Lee, 1999; Tversky, Zacks, Lee & Heiser, 2000), which had already revealed that the arrow triggers dynamic responses and has the capacity to indicate motion and the direction of motion in static visuals for relatively high educated viewers in a Western context. The results obtained in the present study indicate that these results can be generalized to low educated South Africans.

However, one must be careful with jumping to the conclusion that the arrow is an effective indicator of motion and the direction of motion in static visuals for all viewers. The overview of the interpretations of the arrow by the respondents reveals that only half of the respondents is familiar with the arrow as an indicator of (a specific) motion in static visuals. The familiarity with the arrow by these respondents is likely to have caused the effects of the arrow. Some respondents (13.0%) identified the arrow as an indicator of direction. This meaning might seem an unspecified variant of the intended meaning (indicator of direction of motion in the visual). However, these respondents had not mentioned the arrow spontaneously during the interview. Not even when they were asked after their motivation for an expression including the intended motion, which makes their familiarity with the arrow’s meaning as intended in this context highly questionable. Perhaps these respondents recognized the arrow as a sign on the road showing route directions (a purpose for which also in South Africa the arrow is used), and were unable to spontaneously translate this meaning of the arrow to the context of static visuals. An other 17.4% indicated to be unfamiliar with the arrow. It is not surprising that not all respondents are familiar with the arrow. After all, studies by among other Carstens et. al (2006), Dowse & Ehlers (2004) and Foesenek (2006) had already revealed that visual elements representing abstract notions cannot be expected to be familiar to all relative low educated South African viewers. What is also worth noticing is that no attempts by these or other respondents were made to infer the interpretation of the arrow from its analogical residues referred to by Heiser & Tverksy (2006). Apparently, the relation between the arrow as an abstract symbol and its analogical residues as pointed out by Heiser & Tverksy (2006) is not so obvious to the respondents. It appears that either one is familiar with the arrow as an abstract element, or one is not. This implies that the nature of the abstract visual element ‘arrow’ is probably more conventional and arbitrary in the eyes of viewers unfamiliar with it, than predicted by Heiser & Tversky (2006).
2.7.2 Effect of education level and age

Dynamic expressions
Do education level and age influence the type of expressions? The results show that the higher educated a person, the more often he or she spontaneously responded by a dynamic expression to Difficult Visuals. This implies that the arrow triggers more spontaneous dynamic expressions as education level increases. The finding is in accordance with the expectations, as previous studies (e.g. Carstens et al., 2006 and Foesenek, 2006) indicate that the higher a respondent is educated, the more likely he or she is to be familiar with (partly) conventional visual elements representing abstract notions (such as the arrow). Familiarity with the arrow is again likely to result in more dynamic expressions. That this influence of education level was found for Difficult Visuals is also in line with the expectations. The spontaneous dynamic expressions by the respondents in the studies discussed in section 2.1.3 also revealed an influence of the arrow for static visuals displaying relatively unfamiliar fragments (‘Difficult Visuals’) resulting in a relatively great dependence on the presence of an arrow for the recognition of motion. With regard to Easy Visuals it was found that the younger a respondent, the more often he or she responded with a spontaneous dynamic expression. This implies that the arrow triggers more dynamic expressions as age decreases. It appears that Easy Visuals with arrows are not equally ‘easy’ for all respondents. For these visuals it is not the level of formal education, but the time passed since on has been to school that seems to be of influence.

Summarizing, the hypotheses with regard to the factors Education Level and Age can be partly confirmed. Effects of these factors were in particular expected for Difficult Visuals including arrows, which is the indeed the case for Education Level. For Age an effect was found as well, but for Easy Visuals including arrows.

Correct dynamic expressions
No correlation for the visuals with arrows was found between Education Level or Age and the proportions correct dynamic expressions. The supposed influence of Age and Education Level on CD-expression was not revealed. This might seem to be inconsistent with the findings about the influence of Education Level by e.g. Carstens et al. (2006); Dowse & Ehlers (2004) and Foesenek (2006). However, it must be taken into account that the variation of Education Level among the respondents is small (the education level of 78,2% of the respondents was Grade 6, 7 or 8 for both visuals variants; the standard deviation is 1.94 for the +arrow visuals and 1.93 for the –arrow visuals). This makes it difficult to detect a possible influence of Education Level. Nevertheless, other results do imply an influence of Education Level on the interpretation of the arrow as an element indicating the direction of a motion. For the four visuals including the human body (of which a majority of 75% consists of the visuals displaying gymnastic exercises, which are Difficult Visuals), it was found that arrows were more often identified as the source for the correct dynamic expressions (i.e. as an indicator of the direction of motion) as education level increases. In addition, for the visuals displaying objects, it was found that the higher educated a respondent, the more often he or she identified the arrow in combination with hands or an object as the source for the correct dynamic expression. These results
do support the conclusions by Carstens et al. (2006), Dowse & Ehlers (2004) and Foesenek (2006) about the (positive) influence of Education Level on familiarity with and the interpretation of visuals including visual elements representing abstract notions.

### 2.7.3 Other visual elements suggesting motion

As indicated by the results, not all respondents are familiar with the arrow, and if so, it seems that they do not always take the abstract element into account. In addition, its meaning is not spontaneously inferred from its supposed analogical residues in reality. Even when no arrows were present, in many cases respondents responded to Easy Visuals in terms of (the intended) motion. Combining this finding with the frequent identification of visual elements other than the arrow as the source for the intended motion, it seems that other visual elements – analogical, natural ones – are likely to play a role as well when it comes to suggesting (a specific) motion in static visuals.

Analogical, natural elements, especially hands, were identified as the source for the correct dynamic expression to a relatively great extent. Hands, relating closely to the respondent’s bodily experiences, were perfectly recognized by the viewers. When for instance displayed with (an) object(s), hands seem to have a strong capacity to trigger a respondent’s bodily experiences with the elements displayed, possibly consisting of non-static, motion including activities, as was also found by Carstens et al. (2006) and Foesenek (2006). Their results show that actions (and mental states) of humans are more successfully expressed by bodily elements that are familiar on the basis of daily experiences than by applying literacy conventions. Hands seem to suggest the presence of an actor, and manipulation of the object(s) by this actor. If the viewer is also familiar with the object displayed (e.g. how it is supposed to be manipulated or used), a very specific action (and motion) may be triggered by the visual, as “seeing objects leads to detect their affordances, and to potential affordances linked to our previous interaction with objects” (Borghi, 2006), for which empirical evidence was found by among others Tucker & Ellis (e.g. 1998; 2000; 2001; 2004). Carstens et al. (2006) explain the expressive power of the human body by arguing that “all humans have comparable experiences with basic actions and bodily expressions associated with them”. This reasoning is in line with the literature on embodiment and affordances, which offers consistent explanations for the power of displayed hands manipulating objects, triggering familiar (bodily) actions and suggesting motion. Among others Glenberg (2002) argues that the meaning of a situation for an individual is the set of available actions in that situation. According to him, the mesh of affordances, experiences and goals is the basis for this set of actions. Gibson (1979) was the first one to come up with the notion of affordances and he argues that affordances are possibilities for action, determined by the relation between the physical situation and the actor’s body. Borghi (2006) notes that “cognition is embodied, that is, […] it depends on the experiences that result from possessing a body with given physical characteristics and a particular sensory-motor system”, a viewpoint shared by e.g. Barsalou (1999); Berthoz (1997); Cangelosi, Bugman & Borisuyk (2005); Clark, 1997; Prinz, 2002; Lakoff & Johnson, 1999; Langacker, 1986; Hommel, Müsseler, Aschersleben, & Prinz (2001); O’Regan & Noe (2001); Pulvermüller (2003); Rizzolatti & Craighero (1994). For the context of action- and motion-expression in
visuals the theory of embodiment and affordances, in which (bodily) experiences play a crucial role with regard to ascribing meaning to what one perceives, implies that beside the presence of objects, the presence of bodily elements in visuals may help to express action in these visuals as well. The bodily elements, implying the presence of an actor, may be interpreted as explicit indicators for (1) an action being performed with the specific object(s) and (2) the direction and/or goal of the motion which is part of the action. In addition, the presence of bodily elements may influence what specific experiences of the viewer with the object are triggered because the (position of) bodily elements may evoke certain action-experiences one has had with the object (if familiar).

The frequent identification of the hand as the indicator of the intended motion, appears to be consistent not only with the theory of embodiment, but also with earlier empirical studies. Szichcinski (1979, 1980) found that viewers tend to interpret images of bodies even literally. The viewers tried to imitate the hand positions in the visuals when they operated machinery, also when the hand positions were intended to serve as a general guide. Furthermore, in his study into the cognitive processing of images of bodies and hands, Parson (1987) found viewers to imagine themselves in the depicted body positions. While doing so, they took into account real word constraints. For instance, reaction times were longer when the position of the depicted hand considerably differed from their own, when the depicted position was unusual and when it would be unconformable to hold. These results indicate that viewers actively process depictions of body images and that they are likely to take them as guides for their own actions.

Summarizing, in addition to the arrow, the hand also seems to have the expressive power to suggest motion in static visuals as well. This gives rise to a second study, in which it will be examined to what extent the presence of the hand in addition to or instead of the arrow contributes to triggering expressions in terms of (the intended) motion. Contrary to the arrow, for an analogical element such as the hand, no familiarity with conventions is required. In the current and previous studies, hands (and other bodily elements) were perfectly recognized by the viewers. In the context of health education materials, aimed at a broad audience including viewers who are unfamiliar with certain (pictorial) conventions, it would therefore be interesting to know to what extent the hand may serve as support for the arrow in suggesting motion and the direction of motion, or even as an analogical equivalent for the arrow.
3  Study 2: suggesting motion by analogy (hands)

3.1  Introduction

The effect of the arrow has been revealed in Study 1. Of central focus in Study 2 is the supposed analogical variant of the arrow: the hand. The results in Study 1 gave rise to this follow-up study, of which a report is presented this chapter. First of all, in the next section (3.2) the research questions will be outlined. Subsequently, in section 3.3. the method will be presented. The chapter will continue with an overview of the results (section 3.4). In the closing section (3.5.), conclusions will be drawn from those results, and these conclusions will be briefly discussed.

3.2  Research questions and hypotheses

Two sorts of comparisons are made in the second study. To compare the effects of hands with the effect of arrows, expressions triggered by visuals including hands versus visuals including arrows are compared. Furthermore, to examine to what extent hands trigger motion in visuals with arrows, visuals including arrows and hands are compared with visuals including arrows, but with no hands. The first research question is the following:

RESEARCH QUESTION 4: Is the hand an effective analogical equivalent for the arrow in triggering (the intended) motion?

The first hypothesis (applying to both Difficult and Easy Visuals) relating to this research question is based on the observation in Study 1 that all respondents were familiar with hands, which were found to trigger the performance of an action by an actor – and as a result motion – to a great extent, while not all respondents were familiar with the arrow.

HYPOTHESIS 5: The hand is an effective equivalent for the arrow in suggesting motion; displaying hands will lead to more spontaneous D-expressions compared to when only an arrow is displayed.

With regard to suggesting the direction of the motion, the hand displayed in Easy Visuals is expected to yield more CD-expressions compared to the arrow. This is based on the following reason. As was found in Study 1, the actions displayed on Easy Visuals were highly familiar to the respondents. For most respondents, no arrows are required in these visuals to lead him or her to a D- or a CD-expression. Hands on the other hand are likely to have played a significant role in suggesting these familiar actions, as they were identified as the source for the CD-expression relatively often. As a
consequence, the absence of hands is likely to negatively affect the amount of CD-expressions to a greater extent than the absence of the arrow. With regard to Difficult Visuals, the matters are somewhat more complicated. On the one hand, the dependence on the arrow for a CD-expression is relatively high, as the objects displayed in these visuals seem trigger specific actions (and as a consequence specific motions) to a relatively small extent. From this viewpoint one could expect the arrow to be more useful than hands, as the arrow has the power to indicate a specific (that is, the intended) motion. On the other hand, this power of the arrow cannot be expected to affect all respondents, as the results obtained in Study 1 revealed that only 52.2% of the respondents in this study were familiar with the arrow as an indicator of (a specific) motion in static visuals. Furthermore, the absence of hands might make Difficult Visuals even more unfamiliar. Altogether, these phenomena are expected to balance each other out in terms of outcomes. In other words, no substantial differences are expected between the extents to which Difficult Visuals with either arrows or hands trigger CD-expressions. The hypothesis relating to the CD-expressions is as follows:

**HYPOTHESIS 5:** The hand is an effective equivalent for the arrow with regard to suggesting the direction of motion for *Easy Visuals*; for these visuals the spontaneous D-expressions will be correct more often when a hand is displayed compared to when only an arrow is displayed.

Study 2 will also focus on the question to what extent the presence of hands triggers motion and direction of motion:

**RESEARCH QUESTION 5:** Does the presence of hands trigger (the intended) motion in visuals with arrows?

When in addition to the arrow, hands are displayed, an extra source from which motion can be inferred is present. Therefore the combination of hands and arrow is expected to suggest motion to a greater extent compared to when no hands in visuals with arrows are present. The same is expected with regard to suggesting the intended motion. This is based on the assumption that the hand suggests motion by suggesting the performance of an action. The arrow has the capacity to indicate the specific direction of this motion. The interplay of both elements is expected to suggest the intended motion to a greater extent compared to when only one element, the arrow, is displayed. This reasoning results in the following hypothesis (applying to both Difficult as well as Easy Visuals):

**HYPOTHESIS 6:** Visuals with arrows and hands will trigger motion and the intended motion to a greater extent compared to visuals with arrows but without hands; i.e. the presence of both hands and arrows will trigger more D- and CD-expressions compared to when no hands are displayed.

With regard to the influence of Education Level and Age, the expectations are similar as in Study 1 (see also hypotheses 3 and 4 in section 2.2).
Also similar to Study 1, throughout all the results, attention will be paid to the influence of a second response moment (elicitation), providing a second opportunity for the respondent to come up with a response, in which he or she is guided, or directed, towards D- and CD-expressions. However, again, for the same reason as outlined in Study 1, the spontaneous expressions are of greatest interest.

3.3 Method

The method used for Study 2 is highly similar to the one applied in Study 1. For efficiency reasons the focus of this method section will therefore be on the adaptations applied to the method of Study 1.

3.3.1 Respondents

Twenty-three additional respondents were interviewed for the second study. Their details are presented in Table 3.1, together with the details of the two groups of respondents from Study 1.

Table 3.1: distribution of respondents among visual variants; mean grade (education level) and age in relation to Visual Variant (standard deviation between brackets) and percentage of males and females in relation to Visual Variant

<table>
<thead>
<tr>
<th>Visual Variant</th>
<th>n</th>
<th>Grade</th>
<th>Age</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1 + hands/- arrow</td>
<td>23</td>
<td>7.26 (1.94)</td>
<td>32.48 (14.08)</td>
<td>47.8%</td>
<td>52.2%</td>
</tr>
<tr>
<td>Study 1 +hands+ arrow</td>
<td>23</td>
<td>7.48 (1.93)</td>
<td>30.65 (9.31)</td>
<td>43.5%</td>
<td>56.5%</td>
</tr>
<tr>
<td>Study 2 -hands/+arrow</td>
<td>23</td>
<td>7.26 (1.21)</td>
<td>33.50 (13.85)</td>
<td>47.8%</td>
<td>52.2%</td>
</tr>
<tr>
<td>Overall All</td>
<td>69</td>
<td>7.37 (1.91)</td>
<td>31.57 (11.84)</td>
<td>45.7%</td>
<td>54.3%</td>
</tr>
</tbody>
</table>

Similar to the majority of the respondents who participated in Study 1, the 23 additional respondents were all inhabitants of the township Atteridgeville in Pretoria, and rewarded with twenty Rand for their participation.

3.3.2 Research materials

Selection of the research materials

The materials consist of three versions of five visuals. The materials used for Study 2 comprise a selection of five out of the eleven visuals used in Study 1. The main criterion for the selection was that the manipulation – removing the hands from the depiction – would result in a depiction that still makes sense, i.e. that the visual still represents realistic fragments. As a result, the visuals that met the criterion are all visuals in which an object is manipulated by (a) hand(s), given that removing the hands from visuals including the human body only would not results in visuals meeting this criterion. The visuals that did meet the criterion are presented in appendix G. The selection includes two of the
Difficult Visuals (Asthma Pipe; Bottle and Pipe) and three of the Easy Visuals (Pot and Lid; Condom and Package, Cigarettes and Dustbin). All visuals display one or more (parts of) objects, of which one (or a cluster, as is the case for the visuals Cigarettes and Dustbin) is intended to be interpreted as moving into a specific direction.

### 3.3.3 Interview setting

All 23 interviews were held in the same health clinic in Atteridgeville as the interviews in Study 1 and were also recorded by a microphone and a webcam.

### 3.3.4 Procedure

The introduction and briefing of the respondents were similar to those in Study 1. The five remaining visuals (presented one by one, respectively Pot and Lid, Bottle and Pipe, Condom and Package, Asthma Pipe and Cigarettes and Dustbin) were all visuals relating to the domestic environment and were therefore introduced as such. The additional question “What can you learn from this visual?” was left out of the interviews, as this aspect was not of interest in this study. The average duration of an interview was approximately five minutes.

### 3.3.5 Design

The effect of Visual Variant (+hands/+arrow; +hands-arrow; -hands/+arrow) is measured by means of a between subjects design. Furthermore, the experiment uses a repeated measurements design with Response Moment (with levels: spontaneous; spontaneous + elicited) as a within-subject factor, and the proportions D-expressions and CD-expressions (treated as a subset of the D-expressions) as the dependent variables. In order to gain insight into the question to what extent hands in addition to the arrow support the triggering of D-expressions (possibly including the CD-expression), the Visual Variants +hands/+arrow and –hands/+arrow are compared. In order to examine to what extent the hand may function as an analogical equivalent for the arrow, the Visual Variants +hands/-arrow and –hand/+arrow are compared. Of course, in these comparisons only the five visuals selected visuals presented in section 3.3.2 are factored in.

### 3.3.6 Analyses

#### Transcript and Kappa scores

Again, the interviews were transcribed. Similar to the approach in Study 1, a second analyst coded 43% of the responses with regard to the dynamic versus static expression type and the presence versus absence of the intended motion in the D-expression.

The resulting Kappa scores of the two analysts are the following:

1. Static versus Dynamic expression: .88
2. For D-expressions: CD-expression versus no CD-expression: .96
   The majority of differences between the two analyses appeared to be caused by slight structural
differences in the applied definitions of both concepts by the analysts. In a short discussion between
both analysts, all disagreements were resolved.

Coding of the transcripts
   The interviews were transcribed and coded by a first analyst following the same criteria and
   procedures as in Study 1. The coding categories differ with respect to one aspect: the interpretations
   of the arrow. Two additional categories (indicator of focus and spear) could be distinguished from the
   interpretations by the respondents, resulting in the following list of categories:
   - Abstract elements:
     - Indicator of motion, and the direction of this motion in visual
     - Indicator of direction (not further specified)
     - Indicator of focus
   - Analogical element: spear
   - Respondent is unfamiliar with arrow

Statistical analyses
   The statistical analyses applied are similar to the ones in Study 1\textsuperscript{16}. For one respondent the proportion
   CD-expressions could not be calculated for the Difficult Visuals, as this respondent did not respond to
   any of the Difficult Visuals with a D-expression. This was also the case for one of the respondents with
   regard to the Easy Visuals. Therefore, the number of respondents taken into account with regard to
   the proportions D-expression to including the CD-expression, is 22 instead of 23 (for the Difficult as
   well as the Easy Visuals).

3.4 Results
   In this part the results of Study 2 are presented. First of all in section 3.4.1, the focus will be on the
   effect of hands versus the effect of arrows. Secondly, in section 3.4.2 the results regarding hands as
   an addition to arrows will be presented. In section 3.4.3 it will be continued with an overview of the
   various interpretations of the arrow by the respondents in which the interpretations of the arrow in
   Study 1 are completed with the ones yielded by Study 2. Subsequently, in section 3.4.4 the results
   regarding the sources for CD-expressions identified by the respondents are presented. Finally, the
   relation between Education Level as well as Age and the interpretation of the visuals without hands
   will be of central focus in section 3.4.5.

\textsuperscript{16} As was the case for both conditions including hands, no correlation between Age and Education Level was
   revealed by a one-tailed Pearson correlation analysis ($r=.-17$, $p=.44$)
With regard to the interpretation of the results, it is important to take into account that the proportion correct dynamic (CD) expressions is a subset of the proportion dynamic (D) expressions. Each proportion CD-expressions relates to the amount of D-expressions including the correct dynamic expression, i.e. the expressions including the intended motion indicated by the arrow.

### 3.4.1 Hands versus arrows

In order to gain insight into the question to what extent the hand is an effective analogical equivalent for the arrow when it comes to supporting D- and CD-expressions, comparisons were made between the proportions yielded by the visuals with either hands, or arrows.

**Dynamic expressions**

Table 3.2 presents the proportion D-expressions in relation with Level of Difficulty, Response Moment and Visual Variant.

<table>
<thead>
<tr>
<th>Level of Difficulty</th>
<th>Spontaneous</th>
<th>Spontaneous + Elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ hand</td>
<td>+ arrow</td>
</tr>
<tr>
<td></td>
<td>(n=23)</td>
<td>(n=22)</td>
</tr>
<tr>
<td>Difficult (n=2)</td>
<td>.50</td>
<td>.48 ns</td>
</tr>
<tr>
<td>Easy (n=3)</td>
<td>.87</td>
<td>.41 ****</td>
</tr>
</tbody>
</table>

NB: * : p<.05,   ** : p<.025,   *** : p<.01,   ****: p<.001,    ns : not significant

For Difficult visuals, no overall effect for Visual Variant was found (F(1,44)=1.69; p=.11). An overall effect was found for Response Moment (F(1,44)=55.05, p<.001, η²=.56). Split analyses revealed an effect for Visual Variant (F(1,44)=4.57, p<.025, η²=.09) for the spontaneous and elicited responses combined. No interaction was found between Response Moment and Visual Variant (F(1,44)=1.71; p=.09). When in addition to the spontaneous responses, the elicited response are taken into account as well, D-expressions occurred more often for visuals with hands (.94 versus .78), and for the spontaneous and elicited responses combined (.86 versus .49).

For Easy visuals, overall effects were found for Visual Variant: F(1,44)=30.44 p<.001, η²=.41 and Response Moment: F(1,44)=25.14, p<.001, η²=.36. D-expressions occurred more often for visuals with hands (.91 versus .52), and for the spontaneous and elicited responses combined (.79 versus .64). There was an interaction between Visual Variant and Response Moment: F(1,44)=6.90, p<.05, η²=.14). Split analyses showed that the effect of Visual Variant had to be attributed to the spontaneous expressions: F(1,44)=35.76 p<.001, η²=.45, as well as the spontaneous and elicited D-expression combined: F(1,44)=16.79 p<.001, η²=.28).

**Correct Dynamic expressions**

Table 3.3 presents the proportion CD-expressions in relation with Level of Difficulty, Response Moment and Visual Variant.
Table 3.3: Proportion CD-expressions in relation with Level of Difficulty, Response Moment and Visual Variant

<table>
<thead>
<tr>
<th></th>
<th>Spontaneous</th>
<th>Spontaneous + Elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- arrow</td>
<td>+ arrow</td>
</tr>
<tr>
<td></td>
<td>+ hand</td>
<td>- hand</td>
</tr>
<tr>
<td>Difficult (n=2)</td>
<td>.30</td>
<td>.18 ns</td>
</tr>
<tr>
<td>Easy (n=3)</td>
<td>.79</td>
<td>.57 **</td>
</tr>
</tbody>
</table>

NB: * : p<.05,  ** : p<.025,  *** : p<.01,  ****: p<.001,  ns : not significant

For Difficult visuals, overall effects were found for Visual Variant: \( F(1,43)=4.57, p<.025, \eta^2=.10 \) and Response Moment (\( F(1,43)=31.41, p<.001, \eta^2=.42 \)). D-expressions were correct more often for visuals with hands (.49 versus .26), and for the spontaneous and elicited responses combined (.51 versus .24). There was an interaction between Visual Variant and Response Moment: \( F=(1,43)=4.98, p<.025, \eta^2=.10 \). Split analyses showed that the effect of Visual Variant had to be attributed to the spontaneous and elicited D-expressions combined (\( F(1,43)=8.25, p<.01, \eta^2=.16 \); spontaneous expressions: \( F(1,43)=.09, p=.15 \)). For spontaneous responses, it does not matter whether an arrow or a hand is displayed in the visual. When also taking into account the elicited responses however, the non-significant trend in the spontaneous responses showing a higher proportion of CD-expressions among the D-expressions for visuals with hands than with arrows, becomes significant.

For Easy visuals, an overall effect was found for Visual Variant: \( F(1,43)=5.13, p<.025, \eta^2=.11 \) and Response Moment (\( F(1,43)=16.02, p<.001, \eta^2=.27 \). D-expressions were correct more often for visuals with hands (.86 versus .69), and for the spontaneous and elicited responses combined (.86 versus .69). No interaction effect occurred (\( F(1,43)<1.67;p=.10 \)).

3.4.2 Hand and arrow versus arrow only

In order to gain insight into the question to what extent the hand accompanying the arrow supports D- as well as CD-expressions, comparisons were made between the proportions belonging to the visuals with arrows and hands versus the visuals with arrows and without hands. Table 3.4 presents the proportion D-expressions in relation with Level of Difficulty, Response Moment and Visual Variant.

Table 3.4: Proportion D-expressions in relation with Level of Difficulty, Response Moment and Visual Variant

<table>
<thead>
<tr>
<th></th>
<th>Spontaneous</th>
<th>Spontaneous + Elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ arrow</td>
<td>+ arrow</td>
</tr>
<tr>
<td></td>
<td>+ hand</td>
<td>- hand</td>
</tr>
<tr>
<td>Difficult (n=2)</td>
<td>.70</td>
<td>.48 ***</td>
</tr>
<tr>
<td>Easy (n=3)</td>
<td>.86</td>
<td>.41 ****</td>
</tr>
</tbody>
</table>

NB: * : p<.05,  ** : p<.025,  *** : p<.01,  ****: p<.001,  ns : not significant
For Difficult visuals, an overall effect was for Visual Variant ($F(1,44)=9.34, p<.01, \eta^2=.18$) and Response Moment ($F(1,44)=33.00, p<.001, \eta^2=.43$). It appears that D-expressions occurred more often for visuals with hands and arrows (.84 versus .63), and for the spontaneous and elicited responses combined (.88 versus .59). No interaction was found between Response Moment and Visual Variant ($F(1,44)=.05; p=.42$).

For Easy visuals, overall effects were found for and Visual Variant ($F(1,44)=34.33, p<.001, \eta^2=.44$) and Response Moment ($F(1,44)=28.53, p<.001, \eta^2=.40$). D-expressions occurred more often for visuals with hands and arrows (.91 versus .53), and for the spontaneous and elicited responses combined (.79 versus .64). There was an interaction between Visual Variant and Response Moment: ($F(1,44)=4.39, p<.05, \eta^2=.09$). Split analyses showed that the effect of Visual Variant had to be attributed to the spontaneous responses ($F(1,44)=35.96, p<.001, \eta^2=.45$), as well as the spontaneous and elicited dynamic responses combined ($F(1,44)=20.64, p<.001, \eta^2=.42$).

Table 3.5 presents the proportion CD-expressions in relation with Level of Difficulty, Response Moment and Visual Variant.

<table>
<thead>
<tr>
<th>Spontaneous + Elicited</th>
<th>Spontaneous</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ arrow + hand</td>
<td>+ arrow - hand</td>
</tr>
<tr>
<td>Difficult ($n=2$)</td>
<td>.46</td>
<td>.18 ***</td>
</tr>
<tr>
<td>Easy ($n=3$)</td>
<td>.78</td>
<td>.57 *</td>
</tr>
</tbody>
</table>

NB: * : p<.05,  ** : p<.025,  *** : p<.01,  ****: p<.001,  ns : not significant

For Difficult visuals, overall effects were found for Visual Variant ($F(1,43)=94.34, p<.001, \eta^2=.69$) and Response Moment ($F(1,43)=46.19, p<.001, \eta^2=.52$). D-expressions were correct more often for visuals with hands and arrows (.56 versus .26), and for the spontaneous and elicited responses combined (.59 versus .32). There was an interaction between Visual Variant and Response Moment: ($F(1,43)=8.22, p<.01, \eta^2=.16$). Split analyses showed that the effect of Visual Variant had to be attributed to the spontaneous responses ($F(1,43)=6.37, p<.01, \eta^2=.19$) as well as the spontaneous and elicited dynamic responses combined ($F(1,43)=28.17, p<.001, \eta^2=.40$).

For Easy visuals, an overall effect was found for and Visual Variant ($F(1,43)=3.69, p<.05, \eta^2=.08$) and Response Moment ($F(1,43)=16.06, p<.001, \eta^2=.27$). CD-expressions occurred more often among the D-expressions for visuals with hands and arrows (.83 versus .69 versus .69), and for the spontaneous and elicited responses combined (.84 versus .68). No interaction effect occurred ($F(1,43)=<1.67; p=.10$).

Summarizing, it appears that for Easy as well as Difficult visuals, hands accompanying the arrow do not only support D-expressions, but also CD-expressions.
3.4.3 Interpretation of arrow

How did the respondents in Study 2 interpret the arrow? Table 3.6. provides an overview. The first column with results, displays the results for the respondents in Study 2. The second column shows the results for all respondents who were exposed to the arrow, in order to provide a complete overview of all interpretation of the arrow. The table shows for each category the percentage of respondents who interpreted the arrow in this way spontaneously, as well as the final percentage when the explicit question after the meaning of the arrow is taken into account as well.

Table 3.6: Interpretations of the arrow in relation to Response Moment and Visual Variant, presented per interpretation category and in percentages

<table>
<thead>
<tr>
<th>Abstract element</th>
<th>- hand +arrow (n=23)</th>
<th>Overall (-hand/+arrow; +hand/+arrow) (n=46)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spontaneous</td>
<td>Spontaneous + elicited</td>
</tr>
<tr>
<td>Indicator of motion and direction of motion in visual</td>
<td>56.6</td>
<td>78.3</td>
</tr>
<tr>
<td>Indicator of direction (unspecified)</td>
<td>0.0</td>
<td>4.3</td>
</tr>
<tr>
<td>Indicator of focus</td>
<td>0.0</td>
<td>17.4</td>
</tr>
<tr>
<td>Analogical element: spear</td>
<td>17.4</td>
<td>17.4</td>
</tr>
<tr>
<td>Unfamiliar with arrow</td>
<td>NA</td>
<td>4.3</td>
</tr>
<tr>
<td>Missing</td>
<td>0.0</td>
<td>8.7</td>
</tr>
</tbody>
</table>

NB: NA = not applicable

The table shows that in addition to abstract interpretations, by some respondents in Study 2, the arrow was spontaneously interpreted as a visual element of an analogical nature as well, i.e. as a spear. As a response to the visual *Cigarettes and Dustbin*, one respondent for instance remarked that a spear was falling into the dustbin. Furthermore, the arrow was referred to as an indicator of focus, as a sign pointing to (a) certain visual element(s) in order to draw attention to it. “It shows me where I must look”, is the elicited response by one of these respondents. The percentage of respondents familiar with the arrow in the context of motion indication in visuals (56.5%) is comparable to the percentage yielded by the visuals with hands (52.2%, see section 2.4.2). Relatively few respondents exposed to the visuals without hands are unfamiliar with the arrow (4.3%).

Altogether, the results show that approximately half of the respondents is familiar with the meaning of the arrow as an indicator of a specific motion in visuals, i.e. with the meaning of the arrow as intended in this context. An other 17.4%, of the respondents interpreted the arrow as an abstract element as well (after elicitation), but did not ascribe the intended meaning to it. An analogical interpretation of the arrow was provided by 8.7% of the respondents, and 10.9% of indicated to have no idea what the arrow was supposed to represent. If the respondent came up with an interpretation of the arrow spontaneously, it was ascribed either the intended meaning, or an analogical interpretation (i.e. a spear).
3.4.4 Sources for correct dynamic expression

The results in section 3.4.1 and 3.4.2 already showed that when no hands are present, the proportion D-expressions including the CD-expression is significantly lower compared to when hands are present. However, despite the absence of hands, the respondents exposed to these visuals still came up with CD-expressions. With regard to the other visual variants, including hands, the hands were found to play a substantial role in triggering a CD-expression. Given that no hands are present in the visuals with only the arrow, it would be interesting to see on what sources the CD-expression is based when only the object and the arrow remain as possible indicators. Does this for instance lead to a greater reliance on the arrow?

Table 3.7 presents the results with regard to the identified sources for CD-expressions in relation to all three Visual Variants. Given that the visuals with arrows and without hands only concern visuals including objects, the distinction between two subsets of visuals made in Study 1, is no longer relevant. All visuals included in the second study are in the subset of visuals with objects and without the human body. The proportions are of conditional nature as – similar to the approach in Study 1 – they are calculated over the visuals interpreted in terms of CD-expressions only.

Table 3.7: proportion per identified source for CD-expression, in relation to Visual Variant

<table>
<thead>
<tr>
<th>Source Type</th>
<th>-ARROW</th>
<th>+ARROW</th>
<th>+ARROW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract: Arrow</td>
<td>NA</td>
<td>.21</td>
<td>.42</td>
</tr>
<tr>
<td>Abstract + analogical:</td>
<td>NA</td>
<td>.10</td>
<td>NA</td>
</tr>
<tr>
<td>Arrow + Hand(s)</td>
<td>NA</td>
<td>.06</td>
<td>NA</td>
</tr>
<tr>
<td>Arrow + Object(s)</td>
<td>NA</td>
<td>.04</td>
<td>.01</td>
</tr>
<tr>
<td>Analogical:</td>
<td>.78</td>
<td>.56</td>
<td>.44</td>
</tr>
<tr>
<td>Object(s)</td>
<td>.20</td>
<td>.33</td>
<td>.44</td>
</tr>
<tr>
<td>Hand(s)</td>
<td>.54</td>
<td>.20</td>
<td>NA</td>
</tr>
<tr>
<td>Object(s)+Hand(s)</td>
<td>.04</td>
<td>.03</td>
<td>NA</td>
</tr>
<tr>
<td>No source identified</td>
<td>.22</td>
<td>.13</td>
<td>.13</td>
</tr>
</tbody>
</table>

NB: NA = not applicable

The table reveals that the proportion responses in which an arrow was identified as the source, is indeed higher when no hands (and only an arrow) are present, compared to when hands (and an arrow) are present. The depicted objects and the arrow being the only possible indicators for the CD-expression automatically results in a greater role for both elements when it comes to triggering the CD-expression. The proportion responses in which an arrow was identified as the source for the CD-expression shows the highest increase (from .21 to .42; objects: from .33 to .44).

In the following section (3.4.5), the results regarding the influence of Education Level and Age on the proportion D- as well as CD-expressions are presented.

---

17 After 20% of the ID-expressions, the respondent was not asked after his/her motivation, resulting in missing values. The proportions in table 3.7 are based on the remaining 80% of the ID-expressions.
3.4.5 Effect of education level and age on expression type

Dynamic expressions
Table 3.8 presents the correlations between Education Level and the proportion D-expressions, as well as the correlations between Age and the proportion D-expressions, for the visual variant with arrows and without hands.

Table 3.8: correlations between Education Level and proportion D-expressions, as well as Age and proportion D-expressions, in relation to Level of Difficulty and Response Moment (n=23)

<table>
<thead>
<tr>
<th></th>
<th>Spontaneous</th>
<th>Spontaneous + Elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education Level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficult (n=2)</td>
<td>-.10 ns</td>
<td>-.03 ns</td>
</tr>
<tr>
<td>Easy (n=3)</td>
<td>.36 *</td>
<td>.22 ns</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficult (n=2)</td>
<td>-.42 **</td>
<td>-.13 ns</td>
</tr>
<tr>
<td>Easy (n=3)</td>
<td>-.36 *</td>
<td>-.22 ns</td>
</tr>
</tbody>
</table>

NB: * : p<.05, ** : p<.025, *** : p<.01, ****: p<.001, ns : not significant

The analysis reveals a significant relation between Education Level and the proportion spontaneous D-expressions as a response to Easy Visuals. The higher a respondent’s education level, the higher the proportion a D-expressions (other correlations: p=.46 to .16). Age was found to be a predictor for the nature of spontaneous expressions relating to Difficult as well as Easy Visuals. These correlations are negative, showing that the younger a respondent, the higher the proportion D-expressions (other correlations: p=.22 and .28).

Correct Dynamic expressions
Table 3.9. shows the results of the correlation analyses applied to the proportions CD-expressions.

Table 3.9: correlations between Education Level and proportion D-expressions including the CD-expression, as well as Age and proportion D-expressions including the CD-expression, in relation to Level of Difficulty, Visual Variant and Response Moment (n=22)

<table>
<thead>
<tr>
<th></th>
<th>Spontaneous</th>
<th>Spontaneous + Elicited</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education Level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficult (n=2)</td>
<td>.04 ns</td>
<td>.09 ns</td>
</tr>
<tr>
<td>Easy (n=3)</td>
<td>.35 ns</td>
<td>.18 ns</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficult (n=2)</td>
<td>-.41 **</td>
<td>-.47 **</td>
</tr>
<tr>
<td>Easy (n=3)</td>
<td>-.58 ***</td>
<td>-.47 **</td>
</tr>
</tbody>
</table>

NB: * : p<.05, ** : p<.025, *** : p<.01, ****: p<.001, ns : not significant
No significant relations with Education Level were found (p=.42 to .054). Age on the other hand was found to be an explaining factor for the findings with regard to all types of Response Moment and Level of Difficulty. It appears that the younger a respondent is, the more often his or her D-expressions were correct.

3.5 Conclusions and Discussion

In this section conclusions will be drawn from the results presented in section 3.4, and they will be briefly discussed (a more extensive discussion, about the two studies, is outlined in chapter 4). As was the case in Study 1, the focus lies on the spontaneous responses. Worth remarking is that for all results it must be taken into account that though they are based on a substantial number of respondents ($n=23$ per visual variant), they are based on the exposure of these respondents to five visuals only (two Difficult and three Easy Visuals).

3.5.1 Hands instead of arrows

To what extent do hands have the expressive power to serve as an analogical equivalent for the arrow? With regard to Easy Visuals, hands were hypothesised to trigger more (correct) dynamic expressions than arrows. The results support this hypothesis: Easy Visuals including hands triggered more spontaneous dynamic expressions than Easy Visuals including arrows, and these dynamic expressions were correct more often (i.e. included more of the intended motion indicated by the arrow). As hypothesised, with regard to Difficult Visuals, hands and arrows trigger motion and direction of motion to a similar extent. Interestingly, the insignificant trend in the spontaneous expressions suggesting that dynamic expressions are correct more often for visuals with hands than with arrows, is strengthened by the elicitation. The trend becomes significant when the elicited responses are taken into account as well. This implies that even for Difficult Visuals the expressive power of hands as opposed to arrows should not be underestimated.

The results are in line with the finding by e.g. Carstens et al. (2006); Dowse & Ehlers (2004) and Foesenek (2006) that familiar, natural, analogical elements (such as hands) have a greater expressive power for low educated viewers than visual element requiring conventional knowledge. Moreover, the results are in line with the theory of embodiment (e.g. Borghi, 2006; Genberg, 2002) and empirical evidence for this theory (Parson, 1987; Szichcinski, 1979, 1980), according to which bodily experiences play a crucial role with regard to ascribing meaning to and interpreting what one perceives.

3.5.2 Hands in addition to arrows

In accordance with the expectations, the results reveal that the hand in addition to the arrow supports the triggering of motion and the direction of motion. No matter the level of difficulty, visuals with hands as well as arrows, trigger spontaneous more dynamic expressions than visuals with just an arrow, and
when hands are present these dynamic expressions are correct more often. Even when also the responses after the elicitation are taken into account, this is still the case. These findings are in line with the observation in Study 1 that the hand was relatively often identified as the motivation behind a correct dynamic expression. As a consequence, when the hand is no longer displayed, the proportion (correct) dynamic expressions is relatively low.

### 3.5.3 Interpretation of the arrow

Given that the respondents in Study 2 were also exposed to the arrow, the results with regard to the interpretation of the arrow obtained in Study 1 can be extended with the interpretations of the arrow by the respondents in Study 2. As was the case in Study 1, somewhat more than half of the respondents were familiar with the meaning of the arrow as an indicator for motion in static visuals. If a respondent responded by describing the intended motion, the proportion responses in which an arrow was identified as the source was twice as much as for the visuals including hands in addition to arrows (.44 versus .22). However, the relatively low proportion spontaneous dynamic expressions including the intended motion (.18 for the Difficult Visuals and .57 for the Easy Visuals, increasing to respectively .34 and .80 when the elicited responses are taken into account as well) indicate that despite this familiarity with the intended meaning of the arrow in this context, in many cases this familiarity did not lead the respondent to a correct dynamic expression. This suggests that the arrow was not always taken into account by the respondents familiar with the arrow. Worth noticing is also that 17.4% of the respondents exposed to the visuals with arrow and without hands spontaneously referred to the arrow as a natural, analogical element (a spear). By Tversky & Heiser (2006), the spear was identified as one of the analogical residues of the arrow. It was recognized as such by these respondents. However, they did not take the step of translating the analogical meaning of this analogical visual element into an abstract one. The most logical explanation is that even though for none of the visuals the spear is the most obvious object to be displayed, these respondents were simply not aware of the fact that the arrow (or the spear) was not supposed to be interpreted literally. Furthermore, a third interpretation of the arrow as an abstract visual element, namely indicator of focus, was introduced by 17.4% of the respondents in Study 2. Though not the intended interpretation, this interpretation as well as the one in which the arrow is identified as an indicator of motion, makes sense. After all, in other contexts (e.g. indicating direction on the streets, drawing attention to what the viewer is supposed to focus on) the arrow may indeed be used for these goals. Worth noticing is that also by the respondents in Study 2, no spontaneous inferring or reasoning about the possible meaning of the arrow took place. This is inconsistent with the reasoning by Heiser & Tversky (2006), implying that the meaning of the arrow as an abstract element could be inferred from its analogical residues.
3.5.4 Effect of education level and age

Dynamic Expressions
The proportion spontaneous dynamic expressions was expected to increase as Education Level increases and Age decreases. These expectations were confirmed by the correlation analyses, except for the correlation between education level and Difficult Visuals. It appears that for Difficult Visuals the number of formal years of schooling does not play a role, while the time passed since one has had schooling does. For Easy Visuals, both factors were found to be of influence. These results regarding the influence of education level on the responses of the visuals are in line with the findings by Carstens et al. (2006), Dowse & Ehlers (2004) and Foesenek (2006) about the influence of education level on the interpretation of static visuals by South Africans.

Correct Dynamic Expressions
The younger a respondent, the more often his or her dynamic expression was correct. This correlation was found for Easy as well as Difficult Visuals. No influence of education level was revealed by the analyses. As was the case in Study 1, the low variation among the respondents with regard to education level (the education level of 82.6% of the respondents exposed to the visuals without hands and with arrows was Grade 6, 7 or 8; the standard deviation is 1.21) makes it difficult to detect a possible influence of this factor and might be an explanation for the fact that no relations were found.
4 Conclusions and Discussion Study 1 and Study 2

Pictures are worth a thousand words is a commonly used proverb. This can be expected to be more true in a low educated environment where reading skills are relatively low, which makes the use of written language less effective. However, (the few) empirical studies into visual literacy in the South African context have revealed that low education levels make the use of visuals more difficult instead of more easy. At the same time, low education and literacy levels make visuals a relatively crucial message carrier, however limited they are. The frequent use of visuals in (South) African printed health communication materials is in sharp contract with the little knowledge on visual literacy in this context. Therefore, the two studies reported in this thesis were conducted in order to gain further insights into the effective use of static visuals aimed at low educated South African target audiences.

In this chapter, the results obtained by these studies will be discussed in the first section (4.1). Subsequently, the limitations of the studies will be outlined in section 4.2 and in the final section (4.3.) suggestions for future research are proposed.

4.1 General discussion and conclusions

The static nature of visuals in printed materials is in sharp contract with the extensive messages they often intend to communicate. Frequently, designers aim to suggest the presence of motion, an abstract notion that cannot be visualised directly in static visuals. Various conventional visual elements have been invented to visually express abstract notions. With regard to suggesting motion and direction of motion, the arrow is frequently used visual element. In Study 1, it was examined to what extent this abstract visual element indeed has the capacity to suggest motion, and direction of motion, in static visuals. Previously, among others Hegarty et al. (2003), Heiser & Tversky (2006), Szichcinski (1979, 1980), Tversky et al. (2000) and Tverksy & Lee (1999) had already found that the presence of arrows in visuals triggers dynamic responses (i.e. responses including descriptions of motion), revealing the arrow’s expressive power to suggest motion. The visual stimuli in these studies mainly comprised illustrations of relatively complicated mechanical systems and the experiments were conducted in a Western context. Of main interest in Study 1 is to what extent these results can be generalised to visuals displaying less complicated actions, and to lower educated (South African) respondents. The results obtained in Study 1 reveal that the presence of an arrow in static visuals leads to more spontaneous dynamic responses compared to when no arrow is present. Moreover, these dynamic responses more often included the intended motion (i.e. the motion indicated by the arrow in the visuals including this abstract element) when an arrow was present. This implies that in addition to suggesting motion, the arrow also has expressive power to suggest direction of motion for the research population. As expected, these results were found for visuals which did not already trigger (the intended) motion to a great extent even without arrows, i.e. for the visuals referred to as Difficult Visuals in the reports. Also in accordance with the expectations, the elicitation caused an increase in the amount of dynamic expressions. It appears that when the possible presence of motion
is pointed out to respondents, they as yet tend to describe the visuals in terms of (the intended) motion. For the correct dynamic expressions, even this elicitation could not undo the effect of the arrow. Altogether, these results indicate that also for low educated South African audiences, the arrow has the expressive power to suggest motion and direction of motion in static visuals.

However, despite these findings, one should be careful with considering the arrow as an always effective and the ultimate abstract visual element for suggesting motion and direction of motion in static visuals aimed at low educated South Africa audiences. After all, only somewhat more than half of the respondents appeared to be familiar with the meaning of the arrow in this context. Moreover, there were no indications whatsoever that inferences on the possible meaning of the arrow were made by the respondents. It appears that the analogical residues of the arrow identified by e.g. Tversky (2002) and presupposed to possibly reveal the meaning of this symbol, did not lead any of the respondents to the arrow’s (conventional) meaning. Several respondents referred to the arrow as a spear, one of its supposed analogical residues, but did not translate the recognition of this analogical element into a conventionalized meaning of the arrow to express the abstract notion of motion. Apparently, these respondents were not aware that the arrow (or spear) was supposed to be ascribed a non-literal interpretation. Several other respondents (17.4%) were familiar with one of the other conventional meanings of the arrow (i.e. an indicator of focus or direction). These respondents did not mention the arrow spontaneously during the interview and came up with these interpretations when explicitly asked for it. They were aware of the fact that the arrow was not supposed to be interpreted literally as an analogical visual element (a spear), but as a conventional abstract visual element. Despite this awareness, it appeared that even these respondents did not make the translation to the arrow’s meaning as intended in the visuals. These findings imply that the specific meaning of an arrow as an indicator of (a specific) motion cannot be spontaneously and naturally inferred. The existence of natural analogies, and even familiarity with related abstract meanings of the arrow, do not seem to be sufficient to infer the intended meaning of this abstract visual element. It appears that also elements with natural analogies, that are assumed to have an associative relation with their referent, can be difficult to interpret as supposed to when one has not learned their meaning. Though originally the arrow might have been based on the analogical residues as identified by Heiser & Tversky (2006) and Tversky (2002), in reality the meaning of this visual element seems to be of such a conventional nature that is must be learned in order to be applied as intended in the context of motion suggestion in static visuals.

That not all respondents are familiar with and benefit from the arrow to the same extent, is once more indicated by the influence that education level and age were found to have on the expressions types in response to the visuals. Earlier studies into visual literacy in the South African context (see e.g. Carstens et al., 2006; Dowse & Ehlers, 2004; Fosenek, 2006) had already revealed a positive influence of education level on the interpretation of static visuals including (partly) conventional visual elements. Despite the low variation with regard to education level in the sample of the current studies, resulting in small chances to detect a possible influence of this factor, it was taken into account in both Study 1 and 2. Age was taken into account as well, given that higher a respondent’s age, the more time probably has passed since the formal education took place. Moreover, unlike for education level,
the variation in the sample with regard to this factor was high. In Study 1 the amount of spontaneous dynamic responses to Difficult Visuals with arrows was found to increase as education level increased, implying that the higher educated a respondent, the more he or she benefits from the presence of the arrow. With regard to the other visual variant including arrows (the visuals without hands), a relation between education level and Easy Visuals was found. Apparently, these visuals are more ‘easy’ (i.e. trigger more dynamic expressions) as education level increases. These findings provide additional support for the influence of education level on the interpretation of visuals, as earlier revealed by Carstens et al. (2006); Dowse & Ehlers (2004) and Foesenek (2006). The fact that no further relations with education level were revealed, might be explained by the low variation in education level among the respondents. Age was found to be a predictor for spontaneous dynamic and spontaneous correct dynamic responses to various visual types as well (i.e. for the spontaneous dynamic and correct dynamic responses to Easy and Difficult visuals with arrows and without hands, and for the spontaneous dynamic responses to Difficult Visuals with arrows and hands – for the dynamic responses triggered by the visuals with hands and the correct dynamic responses triggered by the visuals without hands, the relation remained even when the elicited responses are taken into account as well). All found correlations were negative, revealing that the older a respondent, the less likely he or she is to respond with a (C)D-expression to the visuals concerned, and the less he or she seems to benefit from the presence of the arrow. The older a respondent, the more time has probably passed since the formal education took place. Therefore the findings with regard to age can be seen as yet an additional indication for the influence of education on the interpretation of visuals. However, it must be noted that this predictor for the influence of education level is less reliable than education level as such. First of all because the reasoning that the older one is, the more time has passed since the formal education took place, is an assumption, which might not be true for all respondents. Secondly, because age is not only likely to relate to the time passed since formal education took place, but also to other factors. For instance, it might be possible that the relatively younger respondents are or were exposed to modern media and/or education materials including visuals (with conventional elements) to a greater extent than the older respondents. As a result of this possible more frequent exposure to visuals, the younger respondents might be more familiar with how to interpret visuals messages (including symbols). Therefore it is questionable to what extent the found influence of age can be directly linked to the influence of education level. Clear however is that education level is of influence with regard to the interpretation of visuals (and conventional visual elements such as the arrow), given despite the low variation in the sample, influences of this factor were detected in both studies.

The results obtained in the present studies once again stress the importance of exploiting the expressive power of familiar analogical elements, e.g. bodily elements, in visuals aimed at low educated (South African) audiences, as was found by e.g. Carstens et al. (2006); Dowse & Ehlers (2004) and Foesenek (2006) as well. Study 1 revealed the expressive power of the arrow to indicate motion and direction of motion, from which one seems to benefit more as one is higher educated. With regard to visuals in which objects are displayed, Study 2 revealed that the dynamic expressions in response to the visuals with arrows are to a great extent based on the presence of hands being displayed with the arrows. Furthermore, hands were even found to trigger spontaneous dynamic and
spontaneous correct dynamic expression to a similar (for Difficult Visuals) or greater (for Easy Visuals) extent than arrows. The hand is a pre-eminent familiar element, involved in many (motion including) activities performed and/or experienced by viewers. Hands appear to suggest the presence of an actor, and when displayed in relation to (an) object(s), they seem to suggest the manipulation of this object by a human.

However, it is perfectly imaginable that one intents to depict a scene including motion in which the presence of motion can simply not be suggested by analogy. For instance because no object manipulation is involved, as is the case for the visuals depicting gymnastic exercises used in Study 1. Even though these visuals include bodily elements (which were perfectly recognised by all respondents), no specific action or goal can be inferred from what is displayed, and no specific experiences are likely to be triggered (which was supported by the results in Study 1). Only knowledge of the capacities of the human body pose some limitations on what motions can possibly be made, but this knowledge still leaves many options open. For such visuals one is dependent on a conventional element such as the arrow in order to suggest (a specific) motion. However, although the arrow resulted in more (correct) dynamic expressions in response to the visuals depicting gymnastic exercises compared to when no arrow is present, these visuals triggered (correct) dynamic expressions to a relatively small extent, even when an arrow was present. If a respondent came up with a correct dynamic expression (i.e. described the intended motion as indicated by the arrow), the arrow was named as the indicator for this motion, but these correct expressions occurred only a few times. This finding once again demonstrates that at least for the population in the current studies, the beneficial effect of the arrow is strengthened by the presence of familiar analogical elements, which in addition to the arrow, might reveal (a specific) motion. When the content of the visual is not familiar to the viewer and/or does not appeal to any familiar aspects or experiences whatsoever, the arrow does not always seem to be the perfect remedy for suggesting motion and direction of motion. This is also demonstrated by Study 2, in which the arrow was in many cases found to be incapable to compensate for the absence of hands. In Study 1, one of the visuals depicting a gymnastic exercise (the visual Exercise – standing) reminded several respondents of the activity of praying, even though these visuals were introduced as gymnastic exercises. Some respondents even explicitly indicated to be surprised about the presence of the arrow pointing out that the hands would be lowered, because according to them this was not what this person was doing. It appears that a great amount of respondents tended to rely on what is logical and most obvious to them on the basis of reality. This may have the upper hand over the arrow, even when the respondent is familiar with the arrow’s meaning in this context. The results yielded by the visuals depicting gymnastic exercises and the visuals without hands raise the question to what extent the arrow would be an effective indicator of (a specific) motion for low educated (South African) viewers when it comes to more complex, unfamiliar depictions, where no or hardly any cues for (the intended) motion are displayed, such as the mechanical systems in the studies by among others Hegarty et al. (2003) and Heiser & Tversky (2006). Unfortunately, the current studies do not provide any decisive answers to this question.

The findings in Study 1 and 2 demonstrating the expressive power of analogical (bodily) elements in visuals for low educated (South African) audiences, are not only in line with earlier studies (e.g.
Carstens et al., 2006; Dowse & Ehlers, 2004; Foesenek, 2006) among comparable target groups. They are also consistent with the theory of embodiment and affordances and empirical evidence for this theory (see e.g. Borghi, 2006; Gibson, 1979; Glenberg, 2002 and Tucker & Ellis, 1998, 2000, 2001, 2004), as the results indicate that respondents assign meaning to what they perceive on the basis of the (bodily) experiences triggered by the visuals. Especially for the visuals in which hands manipulating an object are displayed, the respondents' expressions revealed that in most cases they do not only perceive the visual elements as such, but that the assign meaning to it in terms of (motion including) actions and/or goals corresponding with reality.

In conclusion, the arrow was found to suggest motion and direction of motion in static visuals. This beneficial effect of the arrow increases as education level increases (and as age decreases). However, one should be careful with considering the use of the arrow as the ultimate method for suggesting (a specific) motion in visuals aimed at low educated (South African) audiences, given that only somewhat more than half of the respondents in these studies appeared to be familiar with the meaning of the arrow as a motion indicator in static visuals. Furthermore, even the respondents who are familiar with the arrow did not always seem to take the abstract element into account. While interpreting the visuals, the respondents were found to rely to a great extent on the analogical elements displayed in the visuals. For visuals including objects, the human hand was found to be of significant importance when it comes to suggesting motion and the direction of the motion, supporting previous findings about the expressive power of bodily elements in static visuals.

For designers, these results imply that arrows can be used to indicate motion and direction of motion in visuals aimed at low educated South African audiences, but that it should be taken into account that probably not all viewers are familiar with the meaning of this abstract element in the particular context and that the expressive power of familiar, natural elements such as hands, should be exploited.

The results discussed in this section should be interpreted with caution, due to several limitations of both Study 1 and Study 2. These limitations are discussed in the next section.

4.2 Limitations of the research

Low educated audiences are an unpopular target group for experts in the field of language and communication. This may partly be due to the fact that it is already complicated to study the communicative competence of ‘normal’ language users. Reasons for the unpopularity of low educated audiences as target groups might be that these target groups are too complex or varied, not sufficiently interesting form a societal or financial perspective or that it is complicated to develop reliable methods to gain insight in their communicative skills. Despite possible obstacles, the current two studies have been conducted, as from a theoretical as well as practical perspective it is important to gain insight into the visual literacy of low educated (South African) audiences. Conducting research of an experimental nature among this target group is a challenging endeavour, given that due to the lack of sufficient earlier sound studies it is hard to anticipate on the results and on what would be
effective, appropriate research methods. Furthermore, the research population is often not familiar with experimental settings and corresponding tasks. In addition, the low education levels, often resulting in limited literacy skills, put limitations on the research methods as well as the tasks that can be fulfilled by the respondents. As a consequence, researchers are mainly restricted to oral and/or visual communication methods for acquiring research data. The challenging nature of research among low educated South Africans resulted in several limitations of the current studies, of which some could have been foreseen on forehand, while others were hard to anticipate on.

Various aspects and concepts should have been better controlled. First of all, the number of visuals in Study 1 \((n=11)\) was a very unfortunate choice, as this resulted in an unbalanced design, which again yielded undesired consequences for the possibilities of statistical analyses and comparisons. It was assumed that the visuals would differ with regard to the depiction of familiar elements and/or actions, and therefore the visuals were categorized in two groups (Difficult vs. Easy Visuals) on the basis of the respondents’ spontaneous responses. By using eleven visuals it should have been noticed on forehand that no matter the distribution of the visuals over the levels of difficulty, this number of visuals would never yield a balanced outcome.

Furthermore, to investigate the effect of hands, only five visuals (2 Difficult and 3 Easy Visuals) were included in the research materials. This was due to the post hoc character of Study 2, imposing restrictions on what visuals could be included in the research materials. This was dependent on which visual used in Study 1 were appropriate for the manipulation (removing the hands) required for Study 2. The results with regard to the effect of hands obtained in Study 2 are based on these five visuals only. It is questionable to what extent those results can be generalised to other visuals displaying uncomplicated actions, or to what extent the results are based on certain features of these specific visuals.

Thirdly, the notion of familiarity should have been better controlled. In the current studies it was assumed that visuals displaying familiar elements, scenes and/or actions would trigger (a specific) motion (i.e. dynamic and correct dynamic expressions) to a greater extent compared to visuals displaying less familiar fragments. However, the familiarity of the respondents with what is displayed on the visuals, was not directly measured, it was (indirectly) inferred from the type of expressions triggered by the visuals. If the respondents would have been questioned on their familiarity with what is depicted in the various visuals, more reliable conclusions could have been drawn on the influence of this factor. A material test into the familiarity with the respondents with what is displayed on the various visuals is required in order to draw reliable conclusions with regard to this factor.

An other aspect with regard to the research materials which can be regarded as a limitation, is the use of visuals which were actually used in printed health communication materials. On the one hand in this way it is stayed as close as possible to the visuals messages which are actually aimed at the target group, on the other hand it puts restrictions on what visuals can be selected, decreasing the amount of control one has over the research materials. In search for visuals without any cues for the what exact action was being performed by the actor, an escape to visuals that were not printed in South African health communication brochures (the visual displaying gymnastic exercises) had to be
made in Study 1. This appeared to be a very crucial choice, as the activities displayed on most visuals selected from existing health communication were very obvious to many respondents, even when no arrows were displayed. It would be hard to detect an effect of the arrow with these visuals. Though the visuals displaying gymnastic exercises played a crucial part in revealing the effect of the arrow, the inclusion of these visuals in the research materials, does not seem to be the most optimal choice. The depiction of for example a ball would have been a better choice, as without any arrows, no motion is suggested at all. Moreover, from such a depiction it is completely impossible to infer the presence and/or direction of motion when no arrow is present.

With regard to education level, more decisive conclusions about the influence of this factor on the expression types could probably have been drawn if the respondents in both samples showed a higher variation with regard this factor. In addition, when the variation is high, one is not reduced to the need to rely on other factors to gain insight into the influence of education levels, such as the factor age (which as pointed out before can only be indirectly related to the influence f education level). With the low variation in the current samples, chances for detecting relations were small. However, it is hard to control for this factor when the research takes place in an environment where the researcher is highly dependent on the situation and on others when it comes to recruiting respondents.

Finally, the following should be taken into account with regard to the kind of expressions triggered by the visuals. As mentioned before (and as can be predicted from the theories on embodiment and affordances), visuals depicting familiar elements tend to trigger certain actions, for instance on the basis of the viewer’s experience with what is depicted. When this is the case, respondents may identify this specific action in their expression (e.g. “this person is cooking” for the visual Pot and Lid). Such expressions, often referring to functional acts, describe the visuals on a more general level than just the (spatial) relation between elements (an example of the latter type of expressions is “the lid is going down to the pot”). Often, these expressions on a more general level make perfect sense. However, sometimes these expressions are too unspecific to reveal the information required to know to what extent the visuals triggers the intended motion, in addition to motion. As a result, there might be dynamic expressions that are not considered to be correct (i.e. that are not considered to include the intended motion) whereas the respondent did notice the intended motion. This raises the question to what extent the expressions by the respondents are reliable reflection of what the respondents perceives and what specific motion(s) the visuals triggers. If a follow-up question for further specifications would have been asked, more specific insights could perhaps have been acquired.

4.3 Suggestions for future research

In addition to the limitations of the research outlined above, there are several more interesting aspects where the current studies do not provide insight in that would be interesting to take into account in future research.
First of all, as mentioned before in section 4.1, the current studies do not lead to decisive answers on the question to that extent the found results can be generalised to visuals displaying more complex (motion including) actions. From a theoretical point of view it would be interesting to see if the arrow is effective when no familiar elements at all are present in visuals, and whether more complex motions that have to be suggested by more than one arrow can be effectively suggested by this abstract visual element. Intuitively, the less visual elements other than the arrow triggering (the intended) motion are present, the more the viewer could be expected to rely on the arrow for a correct interpretation of the visuals. However, the results obtained in the current studies with regard to the visuals depicting gymnastic exercises indicate that at least to many low educated (South African) viewers, this reasoning does not seem to apply. More research will be required to gain further insights into to what extent these results can be generalised to other types of visuals and what (cognitive) principles underlie these contra-intuitive findings. From a practical perspective this would be an interesting focal point as well, as it is perfectly imaginable that in practice one aims to communicate a message involving aspects the viewer is not yet familiar with, but that requires the depiction of more complex scenes and actions. Examples of such depictions are mechanical systems as used in the studies by Tversky and colleagues. Such relatively complex systems might be complicated to non-experts. Nevertheless, for designers it would be useful and important to know to what extent the use of arrow is an effective remedy in complex visuals, when motion is hard to suggest by other graphical methods.

Secondly, it would be interesting to gain insight into what the presence of an arrow triggers on a higher level than just the motion of elements in visuals. Several visuals in the current studies were of an instructional nature (e.g. the visual Asthma Pipe). In the present studies, the respondents were not asked after the possible message behind the instructive visuals. It would however be interesting to know to what extent the arrow triggers the instructive nature of instructional visuals. After all, for a complete correct interpretation of such a visual, the viewer does not only have to recognize motion, but also translate this motion into an act that must be performed by him- or herself.

An other enriching addition to the studies reported in his thesis would be to study the comprehension of (motion in) visuals by low educated South African audiences in rural areas. The respondents who participated in the current studies, were al inhabitants of an urban township. Living in this urban environment automatically results in an exposure to visual messages, such as billboards and the media. This exposure might (unconsciously) result in familiarity with visual messages and how they are supposed to be interpreted. This is not (or at least to a much smaller extent) the case for populations living in rural areas. Responses to visuals by these populations may for instance provide more reliable insights into what (abstract) visual elements can be naturally inferred from analogical residues and which ones are of a more conventional nature, as the responses by these populations are much less likely to be influenced by (unconscious) familiarity with visual messages and abstract visual elements through natural and automatic exposure. The inclusion of production tasks (as done before by Tversky & Lee, 1998 and Heiser & Tversky, 2006 in a Western context) in these possible future studies would probably yield enriching insights as well. Such tasks may reveal how abstract notions are represented by respondents who are not or hardly influenced by prior knowledge on
visually representing abstract notions. It would be interesting to see to what extent similar visual basissystematics as in for instance the Western context (revealed by e.g. Tversky, 2005; Tversky et al. 2000) would be used. In this way, further knowledge on what (types of) visuals visual representations are more or less universal and culturally dependent, can be acquired.

In sum, much more additional research is required to enrich the knowledge so far on the visual literacy of various audiences, including low educated audiences. That this is essential with regard to for instance the South African society, with its diverse population in terms of ethnic groups, languages, cultures, literacy and education levels, high HIV/AIDS infection rates and high need for health education, is hopefully once more stressed by this thesis.
Bibliography


## Appendices

### Appendix A  Initial eighteen visuals for Study 1

<table>
<thead>
<tr>
<th>-arrow</th>
<th>+arrow</th>
<th>-arrow</th>
<th>+arrow</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
<td><img src="image16.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
</tr>
<tr>
<td>-arrow</td>
<td>+arrow</td>
<td>-arrow</td>
<td>+arrow</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
<td><img src="image16.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
</tr>
<tr>
<td>-arrow</td>
<td>+arrow</td>
<td>-arrow</td>
<td>+arrow</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td><img src="image1.jpg" alt="Image 1" /></td>
<td><img src="image2.jpg" alt="Image 2" /></td>
<td><img src="image3.jpg" alt="Image 3" /></td>
<td><img src="image4.jpg" alt="Image 4" /></td>
</tr>
<tr>
<td><img src="image5.jpg" alt="Image 5" /></td>
<td><img src="image6.jpg" alt="Image 6" /></td>
<td><img src="image7.jpg" alt="Image 7" /></td>
<td><img src="image8.jpg" alt="Image 8" /></td>
</tr>
<tr>
<td><img src="image9.jpg" alt="Image 9" /></td>
<td><img src="image10.jpg" alt="Image 10" /></td>
<td><img src="image11.jpg" alt="Image 11" /></td>
<td><img src="image12.jpg" alt="Image 12" /></td>
</tr>
</tbody>
</table>
Appendix B: Final selection of visuals for Study 1

<table>
<thead>
<tr>
<th>Easy visuals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pot and Lid</strong></td>
</tr>
<tr>
<td>(domestic environment; instructive)</td>
</tr>
<tr>
<td><img src="image1" alt="Pot and Lid" /></td>
</tr>
<tr>
<td><strong>Salt and Bowl</strong></td>
</tr>
<tr>
<td>(domestic environment; instructive)</td>
</tr>
<tr>
<td><img src="image2" alt="Salt and Bowl" /></td>
</tr>
<tr>
<td><strong>Condom and Package</strong></td>
</tr>
<tr>
<td>(domestic environment; advisory: safe sex)</td>
</tr>
<tr>
<td><img src="image3" alt="Condom and Package" /></td>
</tr>
<tr>
<td><strong>Soap and Bowl</strong></td>
</tr>
<tr>
<td>(domestic environment; advisory: hygiene)</td>
</tr>
<tr>
<td><img src="image4" alt="Soap and Bowl" /></td>
</tr>
<tr>
<td><strong>Fighting Couple</strong></td>
</tr>
<tr>
<td>(domestic environment; advisory: domestic violence)</td>
</tr>
<tr>
<td><img src="image5" alt="Fighting Couple" /></td>
</tr>
<tr>
<td><strong>Cigarettes and Dustbin</strong></td>
</tr>
<tr>
<td>(domestic environment; advisory: quitting smoking)</td>
</tr>
<tr>
<td><img src="image6" alt="Cigarettes and Dustbin" /></td>
</tr>
</tbody>
</table>
**Difficult visuals**

**Exercise – sit up straight**

**Bottle and Pipe**
(domestic environment; instructive)

**Exercise – sit and spread arms**

**Asthma Pipe**
(domestic environment; instructive)

**Exercise – standing**
Appendix C: Pilot interviews and adaptations

After the preselection, the eighteen selected visuals were tested during pilot interviews in The Netherlands. These interviews were of an explorative nature. The aim was for instance to gain insight into the length of the interview, into the effort it required from the respondents and to test whether the visuals and questions would yield useful responses.

The visuals were presented to nine low educated female respondents. The mean number of years of formal schooling they had had was 4.87 years (standard deviation: 2.22). Their mean age was 33.11 (standard deviation: 6.88). The country of origin for one of the respondents was Thailand, another respondent came from Egypt and the remaining seven respondents were born in Turkey. One on one interviews about the visuals were held in a primary school in Tilburg. For two of the Turkish respondents, one of the other Turkish respondents acted as translator. Five respondents were exposed to the visuals without arrows, four other respondents were exposed to the visuals with arrows. For every visual, each respondent was asked what he or she saw on the visual. Several follow-up questions (such as “Can you be more specific?” or “Do you see any movement in the visual?”) were asked in order to see to what extent follow-up questions would lead to an (elicited) enrichment of the response.

The interpretations of the visuals by respondents lead to several insights regarding the (quantity of) materials. First of all, the length of the interviews was considerably long: 20 to 30 minutes. Despite positive reactions of the respondents, who indicated to enjoy the task of talking about the visuals, 20 to 30 minutes during interviews might lead to running the risk of requiring too much effort from the South African respondents. In order to avoid this risk, it was decided to reduce the number of visuals. Visuals showing high similarity with (an) other visual(s) with regard to the depicted action and/or motion were excluded from the selection. This was also the case for visuals deviating from most other visuals in a certain way. For instance regarding the complexity of the intended motion (e.g., the visual representing blood circulation) or with regard to the presence of – when interpreted literally – unrealistic aspects (e.g. the ‘floating’ spoons, the ‘transparent body’ with lounges and with the blood circulation). Furthermore, the visuals representing an action of practical and/or medical relevance were preferred, given the importance of comprehension of such visuals and messages for the viewers ‘in real life’. This resulted in a selection of eight visuals. The interpretation of the eighteen visuals by the respondents revealed that familiarity with certain depicted elements, such as objects and humans, may evoke dynamic interpretations (including the intended motion), even without the arrow. This observation led to the idea of including visuals without any cues for the intended motion, apart from the arrow (if present). Three coloured line drawings depicting a person performing a gymnastic exercise were selected from the internet. On these visuals, a human adopting a certain body posture is depicted. It is expected that apart from knowledge about the capacities of the human body (which still leaves many options open), no cues for a specific movement are present in and/or activated by these visuals. In this way variation among the visuals with regard to the extent to which they are expected to evoke (specific) dynamic interpretations was included. The final selection of eleven visuals can be divided into three types of visuals, presented in the following section.
During a second pilot study, this final selection was presented to two South African respondents, following the procedure outlined in section 2.5.4. The pilot study revealed no obstacles or need for further adaptations and the respondents indicated that the duration of the interview, about fifteen minutes, was perfectly fine.

**Appendix D: Interview setting**

**Location and recording**
Two interviews were held in a student residence in Pretoria. The remaining 48 interviews took place in a health clinic in the township Atteridgeville. The interviews were held in quiet rooms where no interruptions took place. During each interview session the respondent, the interpreter (interviewer) and the researcher were present. All interviews were recorded with a small microphone and a webcam. Each respondent was asked permission for the recordings. One respondent objected to being visually recorded. For this respondent, relevant visual information, such as the use of gestures, was written down by the researcher.

**Interpreters**
To assure that all respondents were able to formulate their answers in their first language, an interpreter held the interviews and translated the answers into English for the researcher. The two interviews at the student room and the University of South Africa were held by an interpreter who was a 21 years old third year law student of the University of Pretoria with Setswana\(^\text{18}\) as his first language and English as a second language. The other interviews were held by a 31 year old voluntary co-worker of the health clinic, also speaking Setswana as a first language and various other African languages and English as second languages. Both interpreters were instructed to hold the interviews in the structured way described in the following paragraphs and to translate the answers as literally as possible. To be sure that these instructions were followed, two interviews of the first interpreter and three randomly selected interviews of the second interpreter were judged on these aspects by a senior lecturer of the Department for African Languages of the University of Pretoria. She concluded that the instructions were perfectly followed by both interpreters. The interpreters were financially rewarded for their efforts.

**Appendix E: Procedure**

**Introduction and briefing of the respondents**
At the start of an interview session, the interpreter and the researcher introduced themselves to the respondent. This took place in Setswana, in case the respondent would not understand any English. Secondly, the respondent was asked after his or her first language in order to determine the language.

\(^{18}\) One of the nine official African languages in South Africa, as well as the main language in Pretoria’s townships
spoken during the interview. Following, the respondent was offered some refreshments and asked for permission for an audiovisual recording of the interview. The introduction continued with a briefing on the purpose and the procedure of the interview. Each respondent was explained that the goal of the interview was to see whether the visuals he or she was going to be shown, would be clear enough to use in public information documents. The respondent was assured that the interview was not a test, but that his or her answers to the questions asked would yield insight into the suitability of the visuals. In this way each respondent was ascribed the role of a helper instead of a person being tested. This approach was taken in order to avoid the respondent feeling ashamed for possible ‘wrong answers’ and to support him or her to tell what came to his or her mind as uninhibited as possible. Furthermore, each respondent was explicitly told that it was important to answer the questions as elaborately as possible. At last, the interviewer asked whether the procedure was clear and whether the respondent was ready to start with the interview.

Additional information about the procedure

The following can be added to the procedure of the interviews as outlined in the section 2.5.4 of the thesis:

- **Recognition of objects**
  In case the respondents did not recognize certain depicted objects, the interpreter identified the object(s). In this way it was made sure that the respondents were not limited by graphical imperfections during their interpretations of the visuals.

- **Personal details**
  After the actual interview, several additional questions concerning personal details of the respondent were asked. Each respondent was asked for his or her age, education level, employment and asked whether he or she had seen these or similar visuals before. At last, the respondent was thanked for his or her effort and time and rewarded with twenty Rand.

**Pilot studies**

The procedure of the interviews was set up and tested by means of two pilot studies. First of all, as mentioned in section 2.5.2, pilot interviews with low educated respondents took place in The Netherlands. As mentioned before, the pilot study revealed that the interviews should not be too lengthy. In addition, it showed that the questions outlined in the previous sections yielded useful answers and that the question “Do you see any movement in the visuals” may lead to an enrichment of the initial response. For the majority of the CD-expressions as a response to the visuals including arrows, it was not clear to what extent the arrow had led the respondent to this interpretation. This is why it was decided to include the additional question “What makes you think that?” in case the intended motion was mentioned, i.e. in case a CD-expression was given. Not all respondents were capable of expressing themselves sufficiently in a second language. Given the importance of knowing as precisely as possible how each respondent interprets the visuals, it was concluded that all

---

19 This was hardly ever the case and if so, it mostly concerned the visual Bottle and Pipe.
respondents should be interviewed in their first language and that an interpreter would be required for the actual study in South Africa.

Furthermore, the pilot study was very useful in various other, more practical, respects as well. First of all, it provided the opportunity to get used to communicating with respondents of a different cultural and educational background, speaking a different language. In addition, it was an excellent opportunity to try out various options for how to cooperate with an interpreter most effectively. The study for instance revealed that the interviews went most fluently when the interpreter knew the procedure and questions by hard and conducted the interview himself.

The second pilot study in South Africa revealed that the final procedure was adequate and that no further adaptations were required.

Appendix F: Analyses
Coding of the data

Static versus Dynamic expression (D-expression)
An expression is considered to be static if the semantic content of the expression does not include any references to motion or actions necessarily including motion. To be more specific, the following two sorts of expressions are considered to be of a static nature:

- No action is described, i.e. the expression does not include any verbs indicating that a person is fulfilling a certain action. The interpretation is restricted to the identification of elements in the visual. An example of such an interpretation for the visual Pot and Lid is “I just see a lid and a pot”.

- The expression does include verbs describing a certain action being fulfilled. However, this action does is not of a dynamic nature, i.e. it does not imply motion of (one of the) elements depicted in the visual. An example of such an interpretation for the visual Exercise – Sit up Straight is “This person is just sitting”, or for the visual Exercise – standing: “This person is standing still”.

An answer is considered to be dynamic when the description includes a description of motion, or actions necessarily including motion. Interpretations with the following nature are considered to be dynamic:

- A literal description of a specific motion (the subject of motion and it’s direction), e.g. for the visual Salt and Bowl: “The salt is falling down into the bowl”.

- A semi-literal description of the motion. The direction is not literally mentioned, but given the position of the visual element(s), the motion and it’s direction can be easily deduced. “The lid is put on the pot”, is an example of such an answer for the visual Lid and Pot. Given that the lid is above the pot, it can be concluded that the lid must move downwards to be put on the pot.

- The reference to a (functional) action which necessarily includes motion. These interpretations include verbs indicating that an action is being fulfilled, but no (semi-)literal reference to a specific
motion included in this action is specified. Examples of such interpretations are “This person is cooking” for the visual Salt and Bowl or “Someone is exercising” for the Exercise visuals.

Presence versus Absence of CD-expression in D-expression

For each D-expression it was analyzed whether it included the intended motion, i.e. whether it concerned a CD-expression. The CD-expressions include the intended motion, i.e. the motion indicated by the arrow in the visuals with arrows. An interpretation was coded as including the intended motion when the movement indicated by the arrow in the visuals with arrows was either verbally described, or demonstrated by means of gestures (the latter occurred only sporadically, and given that the interpretations including the intended motion are a subset of the dynamic interpretations, these gesture-interpretations were ascribed a dynamic nature as well). An example of a D-expression which is considered as a CD-expression for the visual Pot and Lid is for instance “the lid is put on the pot”, while in the D-expressions “this person is cooking” or “the lid is taken from the pot” the intended motion is not described.

Source for CD-expression

The answers to the follow-up question “What makes you think that” in case the respondent responded with a CD-expression were qualitatively analyzed. This resulted in the following categories and subcategories of sources for the intended motion identified by the respondents. (It must be noted that all sorts of combinations of categories occur as well. Some respondents for instance named the object position as well as the arrow as indicators for the intended motion. These combinations were also coded as such.)

- Analogical element:
  - Bodily elements:
    - Hand(s). The hand is identified as the source for the intended motion. For instance just the presence of the hand (“The hand shows me”) or the position of the hand in relation to an object, revealing that the object is manipulated by the hand (“The hand is on the lid” – visual Lid and Pot). Sometimes the hand was mentioned implicitly. This is considered to be the case when a verb necessarily relating to the hand, and implying an action being performed by the hand, is mentioned. For example the verb ‘holding’ (“I can see because he is holding the condom” – visual Condom and Package).
    - Body. For the visuals on which a human body is depicted (the three Exercise visuals and visual Fighting Couple), one of the identified sources for the intended motion is the body posture. For instance for the visual Fighting Couple: “Because I can see the way he just stretches is hand and is holding her, and the woman is trying to protect herself”.
  - Object(s): One or more objects depicted in the visual are identified as sources for the intended motion. Sometimes knowledge about the object was decisive: “I know because I
have an asthma pipe myself” (visual *Asthma Pipe*), or a certain feature of an object: “I can see because the package has been cut” (visual *Condom and Package*).

- **Abstract element: arrow.** The arrow is explicitly identified as the source for the intended motion. (“The arrow shows me”).
- **No source mentioned.** Sometimes respondents experienced difficulties in identifying the source for the intended motion. Their answers were mostly a repetition of an earlier given answer from which no reason for the intended motion can be inferred, or an indication that they could not answer the question about the source for the intended motion.

**Response moment (spontaneous versus after elicitation)**

To gain more insight into to what extent respondents came up with D- and/or CD-expressions spontaneously, the spontaneity of their answers was taken into account. This distinction is important, as the spontaneous answers after the first question reflect the respondents spontaneous, pure, unguided and natural interpretation, without any guidance into a certain (motion-including) direction. The second response moment reveals to what extent guidance may as yet lead to a dynamic interpretation, possibly including the intended motion.

**Interpretation of the arrow**

For each respondent exposed to the visual variant including arrows it was examined how he or she interpreted the arrow (either spontaneously during the interview of after the explicit question). A qualitative analysis of the interpretations revealed that the following interpretation categories can be distinguished:

- Indicator of motion, and the direction of this motion in visual
- Indicator of direction (not further specified)
- Respondent is unfamiliar with arrow

The respondents who assigned the arrow the function of ‘indicator of direction’, were – even after follow-up question for specifications – not able to be more specific on their interpretation.
Appendix G: Selection of visuals for Study 2

### Difficult Visuals

- **Bottle and Pipe**
- **Asthma Pipe**
- **Condom and Package**

### Easy Visuals

- **Cigarettes and Dustbin**
- **Pot and Lid**