A Simple Procedure for Using Vision Impairment Simulators to Assess the Visual Clarity of Product Features

Joy Goodman-Deane, Sam Waller, Katie Cornish, and P. John Clarkson

Engineering Design Centre, University of Cambridge, Cambridge, UK
{jag76,sdw32,klc49,pjc10}@cam.ac.uk

Abstract. Capability loss simulators give designers a brief experience of some of the functional effects of capability loss, thus helping them to understand capability loss better. Wearable simulators, such as vision simulator glasses, can also be worn while using products and prototypes to help identify usability problems. However, this process can be confusing. This paper presents a simple procedure for using vision impairment simulators to assess the visual clarity of product features. The procedure provides clear results that are linked to the numbers of people in the population affected by the issues identified. It was tested with eight accessibility specialists and product developers. Results indicate that they can use this method effectively, and find it useful.

Keywords: Vision loss, Simulation, Inclusive design, Product assessment.

1 Introduction

In order to design inclusively, designers need to understand the challenges faced by people with capability loss. Capability loss simulators are one way to do this. They give designers a brief experience of some of the functional effects of capability loss for themselves [1, 2]. This can be done by wearing equipment that restricts one's motor or sensory capability, such as body suits [3] or pairs of glasses [4]. Software can also be used to show how things might appear to someone with a sensory impairment, e.g. [5].

Capability loss simulators can help designers and other stakeholders to understand capability loss, and to develop empathy with people with impairment. Wearable simulators can also be worn while using products and prototypes to provide insight into the effect of capability loss on product use and to identify usability problems. As such, they are not intended as a replacement for involving users, because they are limited and cannot convey the full experience of someone with an impairment. However, they provide a useful complement to user involvement, providing initial usability feedback before designs are taken to users and helping designers to internalise information gained through other methods.

However, identifying usability problems using simulators can be confusing. Simulators often do not specify what level of impairment is simulated (e.g. [3]), or how

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2014, Part I, LNCS 8513, pp. 43–53, 2014.
© Springer International Publishing Switzerland 2014

many people have that level of impairment. This can make it difficult to determine whether a usability problem will affect many of the target user group or just a few.

This issue has been addressed through work on the Cambridge Simulation Glasses [6]. These vision impairment simulators can be layered to simulate greater levels of impairment. Recent work has calculated the level of visual ability loss simulated by different numbers of glasses. This can be used to estimate how many people would be excluded from using a product based on the number of glasses it is useable with [4]. However, experience with designers indicates that the figures on vision capability and exclusion levels can be confusing to understand, calculate and apply.

This paper presents a simpler procedure for assessing a product using simulator glasses. The procedure gives clear advice on whether the product needs further work or not, depending on whether it passes or fails a standard benchmark. It is described in Section 2, with examples in Section 3. The statistics and calculations behind the procedure are described in Section 4. Results from evaluating the method with accessibility specialists and product developers are given in Section 5.

2 Assessment Method

2.1 Testing the Assessor's Vision

The first part of the assessment method is a very simple test of the assessor's vision. In this paper, "assessor" refers to the person doing the product assessment. The vision test is important because a person's starting vision affects their level of visual ability when wearing the simulator glasses. With the same number of simulator glasses, someone with average starting vision experiences worse visual ability than someone with excellent starting vision [4]. Furthermore, there is some variation in starting vision, even within a group of young designers with "normal" vision. Knowing the assessors' starting vision abilities means that they can each wear the right number of glasses so that they all experience roughly the same level of visual ability.

The test chart used is shown in Figure 1 on the following page. It is attached to a wall at eye level at a distance of 150cm. Assessors wear their normal corrective glasses or contact lenses, but no simulation glasses. They read down the chart to identify the smallest row on which they can read at least 7 letters correctly. The instructions next to this row indicate which sheet to use in the rest of the procedure. An assistant may be needed, to confirm whether assessors have read the letters correctly or not.

The assessment method is not suitable for use by people with poorer levels of vision (worse than 20/28[1]), i.e. people who cannot read the second line on the test chart. This is because the primary aim of the method is to help designers with good levels of vision to understand the needs of people with vision impairments. By focusing on just

[1] In this paper, visual ability is described using Snellen notation (20/X). 20/20 is often described as "normal" visual ability, although many people have better vision than this. Smaller values of X (e.g. 20/16) indicate better visual ability, and larger values indicate worse.

this target group of designers, the assessment procedure can be greatly simplified to use only two very similar flowcharts (see Section 2.2). The simulators themselves can be used by designers with all levels of starting vision [4] but the assessment procedure is more complicated.

The test chart is based on the letter charts used in the Towards Better Design survey [7, 8]. This means that the levels of simulated visual ability can be compared with results from the survey to indicate the numbers of people with those levels of vision, as described in Section 4. The particular levels of vision assessed in the chart were chosen so that the vision levels while wearing the simulator glasses correspond to particular percentages of people in the population (see Section 4.3).

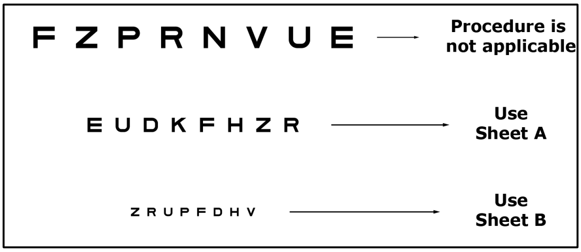

Fig. 1. Vision chart used in testing the assessor's vision (For illustrative purposes only: note that this is not printed at the appropriate size for use in testing vision)

2.2 Assessing the Visual Clarity of a Product Feature

First of all, assessors specify the product they are going to assess. Rather than assessing the whole product at once, they specify a particular user goal with that product, and then assess the clarity of the product features required to achieve that goal. For example, using a mobile phone can involve many different goals, such as charging the phone, making a phone call, and reading a text message. These involve different product features and thus have different levels of accessibility. By examining a single goal at a time, the assessment is more focused and effective. Assessors also specify a use scenario, as levels of accessibility vary depending on aspects such as lighting and the user's prior knowledge.

In the assessment itself, assessors use assessment sheet A or B, depending on their starting vision. The sheets each contain a flowchart, summarizing the procedure to follow, as shown in Figures 2 and 3 on the following page. The flowcharts (and thus the procedure) are identical except for the number of glasses to wear at each stage. Assessors with excellent vision wear more simulator glasses than those with good/average vision to achieve the same levels of simulated vision ability.

Assessors start by wearing the number of simulator glasses listed in the first diamond. The glasses used are the Cambridge Simulation Glasses (Figure 4). These restrict the ability to see fine detail and perceive contrast differences, and can be layered to simulate greater levels of impairment [6].

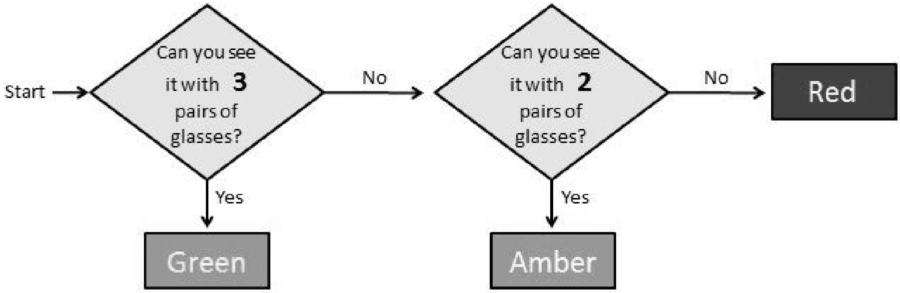

Fig. 2. The assessment procedure flowchart for assessors with good/average vision (visual ability between 20/16 and 20/28)

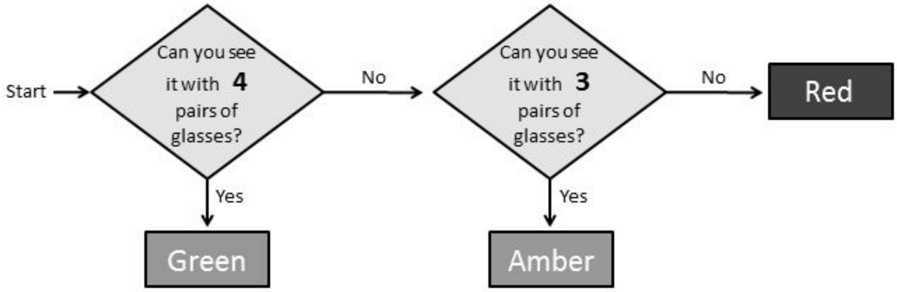

Fig. 3. The assessment procedure flowchart for assessors with excellent vision (visual ability better than 20/16)

Fig. 4. Cambridge Simulation Glasses

The assessors examine the product features while wearing the specified number of glasses. This may involve trying to read text on the product, discern icons, perceive salient details in images or distinguish parts of a product from each other visually. It is important that assessors keep the product at a normal working distance while doing this, to avoid skewing the results.

If the assessor can see the product feature successfully, it is rated "Green" as indicated by the leftmost box in the flowchart. This indicates that less than 1% of the population is excluded from using that product feature on visual grounds (see Section 4 for calculations). The product has passed this basic visual accessibility check, although this does not, of course, guarantee that there are no accessibility problems overall. Ideally, this procedure should be used together with user involvement as part of an overall inclusive design process [9].

If the assessors cannot distinguish the product feature successfully, they remove one set of glasses. Thus people with good vision wear two pairs of glasses, and those with excellent vision wear three, as shown in the second diamond in the flow charts. They then examine the product feature again. If they can distinguish it successfully this time, it is rated "Amber". This indicates that between 1 and 6% of the population is excluded from using that product feature on visual grounds. The design of the feature should be improved if possible, but it may not be the highest priority.

If the assessors cannot distinguish the product feature successfully while wearing the second set of glasses, then it is rated "Red". This means that more than 6% of the population is excluded from using it on visual grounds. It really needs to be improved to reduce user frustration, avoid losing customers, and include the senior market.

3 Example Assessments

Before examining any products, the assessor's vision was tested using the test chart in Fig 1. She could read all the letters on the bottom line of this chart, and thus used the flowchart in Fig 3 for the assessments below.

3.1 Assessment 1: Box of Biscuits

The assessor examined the visual clarity of the ingredients list on a box of biscuits, shown in Fig 5. The product goal was to determine whether the biscuits contained milk. The scenario involved the biscuits being selected by someone unfamiliar with the biscuits or their ingredients, in an indoor setting with good daytime lighting.

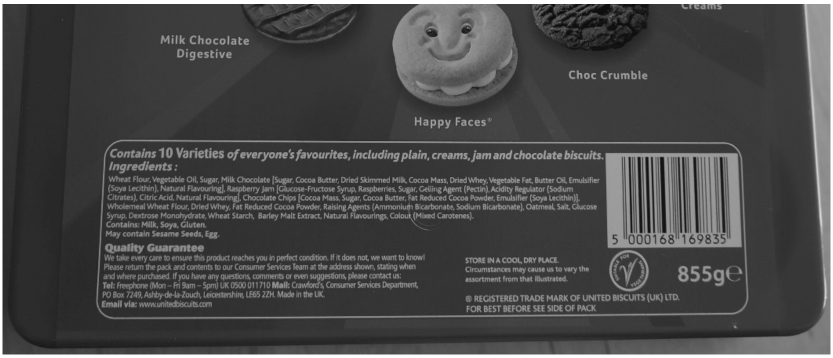

Fig. 5. Portion of the biscuit box used in the assessment, showing the ingredients list

Following the procedure in Fig 3, the assessor started with four pairs of simulator glasses. With these, she could not read the ingredients sufficiently clearly to tell if they contained milk. Therefore, she tried again with three pairs of glasses, as indicated in the second diamond in Fig 3. This time, she could read the ingredients successfully. Thus, this product feature was rated Amber. This means that between 1% and 6% of the population would be excluded from reading the ingredients list. The list should be improved if possible. An obvious way to do this is by increasing the font size. However, there are other ways of doing so without taking up more space on the packaging, such as using a bolder font, or darkening the background to increase the contrast of the letters.

3.2 Assessment 2: Digital Camera

The second assessment examined the visual clarity of the PC/AV port on a digital camera (see Fig 6). The product use task was to insert the cable into the port, in good lighting, without knowing the location of the port on the camera.

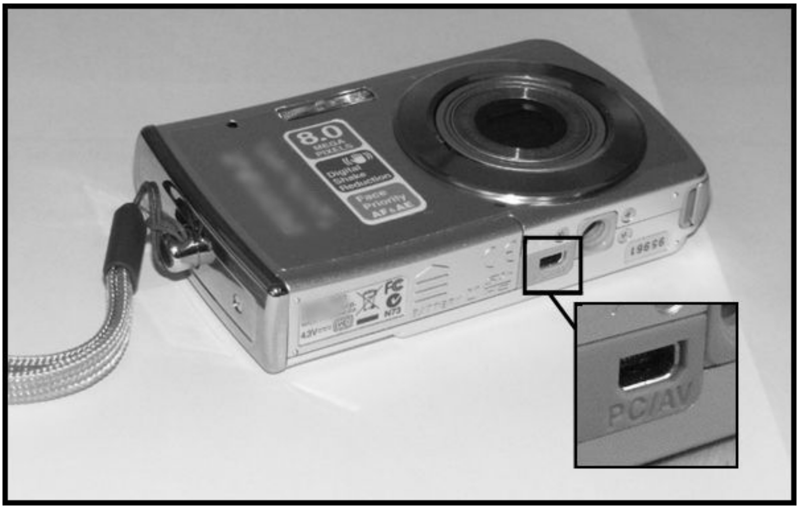

Fig. 6. Digital camera used in the assessment, with close-up of the PC/AV slot

As before, the assessor started with four pairs of glasses. While wearing these, she was able to find a slot of approximately the right size and shape. By trying to insert the cable in this slot, she determined that it was indeed the right slot, and succeeded at the task. Thus the slot was rated Green, as shown in Fig 3. Less than 1% of the population would be excluded from using it on visual grounds alone.

However, users with lower levels of prior knowledge or confidence may not be willing to insert a cable into an unknown slot. Furthermore, many could not rely on touch feedback to determine if the cable fits properly. It is, therefore, also important to examine the visual clarity of the text label next to the slot, which reads “PC/AV”. The assessor was unable to read this while wearing four pairs of glasses, or even with three pairs of glasses. The text label is therefore rated Red, meaning that more than 6% of the population would not be able to read it. It fails our benchmark, and needs to be improved. This could be done by increasing the contrast of the text. To avoid the cost of having to print an additional label, the text could be printed on an existing nearby label (such as that containing the serial number), which could be repositioned appropriately.

4 Underlying Calculations

4.1 Exclusion Calculations

Product features are rated as Green, Amber or Red, depending on the proportion of the population that would be unable to see them, and thus be excluded from their use. The exclusion calculations are best explained using an example, summarized in Table 1. Consider the assessor from Section 3. She could read the bottom line on the vision chart and thus has visual ability better than 20/16. The best human visual ability is around 20/9 [12]. However, this is rare, and the distribution of visual ability in the population means that most assessors in this range actually have vision between 20/16 and 20/11.

Table 1. Exclusion calculations for an example (Amber) assessment

	Example assessor with "excellent" vision (i.e. who can read the bottom line on the test chart)	
	Worst vision in this range	Good vision in this range (most assessors in this range have vision worse than this)
Starting visual ability (VA) (Snellen visual acuity)	20/16	20/11
VA with 3 pairs of simulator glasses	20/49	20/34
Estimated exclusion for a product that is just viewable with 3 pairs of glasses (weighted percentage of survey sample with VA worse than that simulated by 3 pairs)	1.0%	4.9%

We measured twenty participants to determine the effect of different numbers of simulator glasses on people's vision (unpublished study, c.f. [4]). There is some variation between individuals, but it is small enough that the mean values can be used to estimate the assessor's visual ability while wearing three (or four) pairs of glasses (see Table 1).

The assessor could read the ingredients list on the box of biscuits with three pairs of glasses, but not with four (Section 3.2). It is reasonable to assume that people with visual ability (VA) equivalent to (or worse than) four pairs of glasses would not be able to read the ingredients list. They would thus be excluded from it. Similarly, we assume that people with VA equivalent to (or better than) three pairs would be able to read the ingredients and thus not be excluded.

The precise level of VA required to read the ingredients is unknown but lies somewhere between these two values. In the worst-case (most exclusive) scenario, it lies just below the VA with three pairs of glasses. People who have VA equivalent to three glasses can just manage to read the ingredients, and so are included. Everyone with worse VA is excluded. This forms a *conservative* estimate of exclusion. Using this estimate ensures that anyone who is actually excluded is counted in the estimate, at the risk of counting some people who may not actually be excluded (but may still find the product difficult to use). Using a conservative estimate encourages designers to think inclusively.

Thus exclusion is estimated by adding up all the people in the dataset with vision worse than that simulated by three pairs of glasses (see Table 1 for results). More details of how this is done can be found in [4]. The dataset used comes from the Towards Better Design survey [7, 8]. This survey tested a range of user capabilities related to product design. It was chosen because it covers a wide range of vision and other capability measures, meaning that exclusion for other impairments and combinations of impairments can also be calculated [10, 11]. However, in this paper, we focus only on the vision results.

Although the survey was relatively small (362 people), it was postcode sampled across England and Wales and weighted for age and gender. Thus, it can give a good indication of exclusion on a wider scale.

In the example, it is likely that between 1.0% and 4.9% of the population would be excluded (see Table 1). The ingredients list is thus rated Amber (1-6% exclusion). Similar calculations were done for assessors with different starting vision, and different numbers of simulator glasses.

4.2 Choice of Assessment Bands

The assessment procedure gives banded results (Green, Amber and Red) rather than exact exclusion values. This was chosen for two main reasons. Firstly, it is of little interest to a designer whether the exclusion is exactly 2.6%. It is more practically useful to know whether it falls within a band that requires further action (e.g. between 1% and 6%). Secondly, the exact level of exclusion depends on the exact visual ability of the assessor. However, in practice, it is difficult to measure visual ability exactly. For practical reasons, the assessment uses a simplified vision test to estimate the assessor's visual ability within a range (e.g. between 20/16 and 20/28).

The Green, Amber and Red bands correspond to: less than 1%; 1% to 6%; and greater than 6% exclusion. These were chosen because they roughly correspond to levels of exclusion that designers can relate to. Ergonomics books commonly refer to

95th and 99th percentiles, corresponding to 5% and 1% exclusion (e.g. [13], p28). Thus, designers are unlikely to be interested in levels of exclusion below 1%, while values above 5% are high enough for them to realise that something really is remiss.

However, if wearing three pairs of simulator glasses corresponds to about 1% exclusion, then removing one pair increases the exclusion value to about 6%, not 5%. Thus the boundary for the Red band was set at 6% instead of 5%. Ideally, simulator glasses would be chosen that alter exclusion by about 5%. However, the filters for such glasses are not commercially available. The Cambridge Simulator Glasses use the smallest gradation in filters that are easily obtainable commercially.

4.3 Test of Starting Vision

Assessors start the assessment procedure by taking a simple vision test (see Section 2.1). To make this test as simple as possible, different levels of starting vision were combined into ranges, e.g. 20/16 and 20/12 were placed in the same range. The ranges were set so that the exclusion results for people in the same range corresponded to the same band: under 1% (Green), 1-6% (Amber) or over 6% (Red) exclusion. Two people within the same starting vision range should get the same banded result (Green, Amber or Red) with the same number of simulator glasses.

For people with fairly “normal” vision (better than 20/28), only two vision ranges were needed (good/average vision; and excellent vision). Thus only two rows were really needed on the test chart in Figure 1. An extra row was added at the top, to help avoid confusion for people whose vision is worse than 20/28. It also helps people get accustomed to the test chart procedure.

5 Testing the Method

5.1 Workshop

The method was tested in a workshop with eight accessibility specialists and product developers. Level of professional experience varied from 10 months to 15 years. The workshop started with an introduction to the simulators. After this, the starting vision of each person was measured, and they were given the appropriate assessment sheet, as described in Section 2. The assessment method was then explained to them with an example.

The group was then split into pairs. Each pair worked together to assess a variety of items in the meeting room. Choices included TV remote controls, mobile phones, a fire extinguisher, an air conditioning control unit, a packet of adhesive tack, and a microwave. The members of each pair took it in turns to carry out the assessment, while the other person recorded the results. After completing the assessments, the workshop attendees completed a short questionnaire about their experiences of and opinions of the method.

5.2 Results

The results from this initial workshop indicate that participants were able to understand and use the assessment method effectively. However, some clarifications may be helpful to further improve the method. In particular, it would be helpful to provide more worked examples, covering the full range of Green, Amber and Red outcomes, and clarifying how borderline cases should be assessed (where text is difficult but not impossible to see).

All attendees said that they would like to use the simulators again. Seven of the eight said that they would use them to assess visual clarity, with several describing it as a useful technique. For example, one person described it as a *"very useful and consistent check"*.

6 Conclusions and Future Work

This paper has presented a simple procedure for using vision impairment simulators to assess the visual clarity of product features. The procedure gives clear advice on whether the product needs further work or not, depending on whether it passes or fails a standard benchmark. Assessors wear two different sets of simulator glasses while trying to use the product feature. Depending on whether they are successful or not in using the feature, it is rated Green, Amber or Red. These bands reflect the numbers of people who would be excluded from that product feature, and thus the importance of action in improving it.

Participants in a workshop could understand and use the method effectively. Further work is required to test the effectiveness of alternative approaches to resolving the issue of borderline assessment cases, and presenting results in terms of difficulty and/or exclusion. Further testing is also needed with a wider range of industry professions.

Acknowledgements. This work was funded by EPSRC through the KT-EQUAL programme. We would also like to thank the RNIB for helping to organise the workshop where the method was tested. Thanks are also due to the members of the 'Design and Usability Community' group who took part in the workshop.

References

1. Nicolle, C.A., Maguire, M.: Empathic Modelling in Teaching Design for All. In: Stephanidis, C. (ed.) International Conference on Human-Computer Interaction; Universal Access in HCI: Inclusive Design in the Information Society, vol. 4, pp. 143–147 (2003)
2. Cardoso, C., Clarkson, P.J.: Impairing Designers: Using Calibrated Physical Restrainers to Empathise with Users. In: Kose, S. (ed.) 2nd International Conference for Universal Design. International Association for Universal Design, Kyoto (2006)

3. Hitchcock, D.R., Lockyer, S., Cook, S., Quigley, C.: Third Age Usability and Safety - An Ergonomics Contribution to Design. International Journal of Human-Computer Studies 55(4), 635–643 (2001)
4. Goodman-Deane, J., Waller, S., Collins, A.-C., Clarkson, J.: Simulating Vision Loss: What Levels of Impairment are Actually Represented? In: Anderson, M. (ed.) Contemporary Ergonomics and Human Factors 2013, Institute of Ergonomics & Human Factors (2013)
5. Goodman-Deane, J., Waller, S., Clarkson, J.: Simulating Impairment. In: Bryan-Kinns, N., Lloyd, T., Sheridan, J.G. (eds.) (re)Actor3/HCI 2008 (2008)
6. Engineering Design Centre: Inclusive Design Tools. In: Inclusive Design Toolkit, http://www.inclusivedesigntoolkit.com/betterdesign2/inclusivetools/inclusivedesigntools.html
7. Clarkson, P.J., Huppert, F.A., Tenneti, R., Waller, S., Goodman-Deane, J., Langdon, P., Myerson, J., Nicolle, C.: Towards Better Design, 2010 [computer file]. Colchester, Essex: UK Data Archive [distributor]. SN: 6997 (2012), http://dx.doi.org/10.5255/UKDA-SN-6997-1
8. Tenneti, R., Langdon, P., Waller, S., Goodman-Deane, J., Ruggeri, K., Clarkson, P.J., Huppert, F.A.: Design and Delivery of a National Pilot Survey of Capabilities. Applied Ergonomics (2014)
9. Engineering Design Centre: Inclusive Design Toolkit, http://www.inclusivedesigntoolkit.com/
10. Keates, S., Clarkson, J.: Countering Design Exclusion: An Introduction to Inclusive Design. Springer, London (2003)
11. Goodman-Deane, J., Waller, S., Williams, E., Langdon, P., Clarkson, P.J.: Estimating Exclusion: A Tool to Help Designers. In: Proceedings of Include 2011. Royal College of Art, London (2011)
12. Kirschen, D.G., Laby, D.M.: Sports Vision Testing: An Innovative Approach To Increase Revenues. Optometric Management (May 2006), http://www.optometricmanagement.com/articleviewer.aspx?articleid=71635
13. Pheasant, S.: Bodyspace, 2nd edn. Taylor & Francis (1999)