Age-Related Accessibility Biases in Pass-Face Recognition

Ray Adams[1], Gisela Susanne Bahr[2], and Ejder Sevgen Raif[3]

[1] Centre Head
Collaborative International Research Centre for Universal Access (CIRCUA)
School of Engineering & Information Sciences, Middlesex University
The Burroughs, Hendon, London NW4 4BT, UK
[2] Centre Head (USA)
Florida Institute of Technology
150 West University Boulevard, Melbourne FL 32901-6975, USA
[3] Collaborative International Research Centre for Universal Access (CIRCUA)
adams@churchillians.net, e.raif@mdx.ac.uk, gbahr@fit.edu

Abstract. Accessibility and security are often depicted as conflicting aspirations. Accessible systems may be less secure and secure systems may be less accessible. The search is on for greater security for logging onto systems, whilst achieving acceptably high levels of accessibility. Pass-faces are based on the twin axioms of greater accessibility and security. A new user of a pass face system is asked to select "n" faces from an array of faces, where n is at least two and usually more. The user is required to memorize those faces and to recognize them again when represented to you as part of larger display. It has been suggested that this approach is less susceptible to poaching than are alphanumeric methods. There has been a considerable volume of work to evaluate the usage of pass face systems, but little work on the psychology of pass faces. Equally, pass face systems have received little attention from researchers in accessibility. In the present study, two previously unrelated themes were investigated in two experiments. First, are pass face systems acceptably usable? Second, how do pass face systems rely on the reliability of human face recognition memory? In two experiments, two types of pass face system consisting of (a) older faces; over 50 years of age and (b) younger faces; under 30 years of age were created. It turns out that younger participants are often better at recognizing younger faces than older faces in the context of pass face security, whilst older participants are sometimes better at recognizing older faces than younger faces in the context of pass face security. Thus an experiment that used only younger faces would falsely conclude that younger participants were better at face recognition memory than older participants in general. These results were checked and confirmed by literatures reviews of pass face security and human recognition memory for faces. These results show that universal access cannot be applied on a one-size-fits-all basis. They also suggest that the security-related disciplines of HCI and psychology would benefit from greater interaction between them.

1 Introduction

This research tackles two prominently held concerns that are very important for universal access. The first concern is that accessibility and security interests are in

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2011, LNCS 6766, pp. 423–431, 2011.

conflict, the so-called "Security vs. Access Paradox". On one hand, a website or an application must be accessible to intended users, but, on the other hand, adequate security is vital. The World Wide Web is designed to be accessible, but does this focus on the user automatically support security? It is been argued that such a focus can lead to less security. For example, a truly secure browser would request approval for every action, from launching popup window to executing a small script. It would wait for a reply before acting. This will never happen, of course, because most users would simply turn off a system or choose to ignore it, when it asked for a continual stream of approval responses [1]. A shared theme of both accessibility and security is the safe and accurate identification of an intending user. In this study, the use of pass face security systems to support effective user identification is explored. (Such systems are often called "pass-face" systems as they employ faces rather than words or numbers).

The second concern is that older adults are often likely to perform less well than younger adults at cognitive tasks. There is a vast literature that demonstrates cognitive decline through the lifespan (see reference list below for a small sample). Experts tell us that aging leads to "normal" declines in cognitive effectiveness. In particular, such changes include reduced processing speed, greater tendency distractibility plus reduced capacity to process and remember new information i.e. working memory." Such changes can create negative stereotypes about older adults, though they are to be expected and are not symptoms of clinical, cognitive impairment. If so, the rapid presentation of materials, frequent focus changes combined with distractions and intrusions may reduce understanding and recall of information by older adults. In contrast, face recognition is purported to demonstrate an own age bias. If so, older adults may actually be better recognizing older faces from memory than their younger colleagues do. Or at least they may be better working with older faces than younger faces. If such results hold, they would provide important counter examples against the age as total decline viewpoint.

So, here these two concerns about (a) the "Security vs. Access Paradox" and (b) the total age-related decline view, through investigations of pass face security systems are addressed. In so doing, we attempt to bring together two previously unrelated and important themes that both concern pass face systems. Such systems rely on users deploying their recognition memory skills to recognize previously selected faces from amongst presentations of larger sets of faces. The correct selection of these faces acts as the equivalent of the correct selection of a pass word or pin number from a conventional display. First, the user experience of using simple pass face security systems is investigated. How well do users readily learn to use such systems when they encounter them for the first time? Second, since there is research evidence that the efficacy of human recognition memory is influenced by the age of the user and of the faces used, namely an own-age bias, face recognition memory too is explored in this context. Since pass face systems rely on human recognition, then it follows that older users might be expected to perform better with older faces (i.e. of their own age group), whilst younger user might show the opposite preference, being better with younger faces (i.e. of their own age group). To test these predictions, a simple experiment was performed, combining younger and older users with younger and older faces in a pass face system.

It is generally agreed that older adults often perform more poorly than younger adults on a range of cognitive tasks involving perception, memory (recall or recognition) and problem solving. It is said that tasks that require experience and maturity are the only exceptions to this generalisation to be found. If so, then exceptions to this rule are important theoretically and may have important practical implications.

It now turns out that there is some evidence of superior performance by older adults on a specific memory task, including, but not limited to recognition memory for faces. Older adults are reported to be better than younger adults for recognition memory for older faces. In contrast, younger adults are reported to be better than older adults for recognition memory for younger faces. One explanation offered is that people are better at recognition memory for faces for groups with which they can identify e.g. older people with older adults, younger people with younger people etc and that, perhaps, they process such information more thoroughly. The data suggest that this might be an age-specific result.

The focus of the present work is also on some important, practical implications of these results, namely to create a log-in function that is based on recognition memory for older faces, in order to provide older adults with a system that they will find more accessible cognitively.

2 Experiment One

2.1 Introduction

The first study involved the creation of two simple, pass face prototype systems, one using younger faces and one using older faces. The aims were two fold: The first aim was to see if our prototypes were usable, accessible and acceptable to our users, both older and younger. The second aim was to look for an own-age bias in face recognition memory using our pass face prototype systems. The prediction was that older adults would do better than younger adults for recognition memory for older faces. Conversely, younger adults would be expected to do better than older adults for recognition memory for younger faces. In this first study, the focus was on investigating the user experience, ease of use of a pass-face prototype and the learning curve associated with learning to use simple, pass-face systems.

2.2 Participants

The participants were in two samples, (a) older adults (over 60 years of age) and (b) younger adults (under 30 years of age). There were seven participants in each group. The age of each person was recorded but their identities and other demographic details were not (e.g. occupation, location etc) in order to protect their anonymity. Participants were recruited from local sources, subject to the following conditions: No copyright issues were involved and no famous people, no people judged to be unique, very unusual or very easy to recognise were included.

2.3 Materials

The pass-face systems were based on arrays of nine faces at a time, on the screen of a standard laptop, based on simple word processed documents, with included hyperlinks for target words and faces. Half the arrays used younger faces (under 30 years) and the other half were based on older faces (over 60 years of age). Age groups were not mixed within arrays. For each prototype, a pass-face key (a pass-face equivalent to a password) was created. To form the pass-face keys, each time two faces were chosen at random from the sample of younger faces and two faces were chosen each time at random from the sample of younger faces. The pass-face sequence alone was presented before each prototype. Subsequently, a random array of nine faces was presented on one screen, within which the pass-face items were included at random locations. The participant was to select one key face at each presentation. The background items were selected at random from the sample of faces of the correct age group, excluding the current pass-face sequence. Next, a second array of nine faces was presented on one screen, within which the pass-face items were included at random locations. Each face was linked by a hyperlink to an appropriate response page that provided feedback to the participant.

2.4 Methods

Each participant was tested alone. They were each shown a pass-face key of two faces for five minutes and asked to study the faces so that they would be able to recognise them in a bigger set of faces (nine at a time). They were then shown a display of nine faces that included the pass-face sequence items at randomly chosen locations. The participant then selected any one of the two pass-face key items by cursor clicks. As each face depiction was a hyperlink, the act of selection of a face took the participant to a new screen. If their response had been correct, the new screen would contain the message “Well done, that was correct. Now select another one of your faces” and an array of nine faces would be displayed. This process continued until the participant correctly selected both pass-face key items. At this point, the participant was able to a simple Internet based selection task where they could select screens that reflected their hobbies and interests. If the wrong face was selected, the participant was redirected to a different screen that simply said “Apologies, you will have to start again.”

At the start, each participant was given a practice session in which an incorrect response took them back to the initial pass-face sequence on a new screen that contained the message “Apologies, you will have to start again.” To simplify this first experiment, order of testing was held constant for all fourteen participants. Every participant, worked first with five trials with older faces, followed by five trials with younger faces. This allows an initial evaluation of the use of pass-face systems, some indication of own-face recognition bias and the learning curve for such systems. The instructions emphasised accurate performance, asking participants to work as quickly as possible without making errors.

2.5 Results

The time per trial and errors were recorded. The resulting time data are summarised in table one above. An analysis of these data suggests that older faces take significantly longer to use than the younger faces ($F = 16.74$, df 1,68, $p<0.0001$), though this effect could be due to the confounded influence of practice. The two age groups did not differ significantly in overall time to complete tasks ($F = 1.6$, df 1,68, n.s.). There was some indication of an interaction between age of participant and age of face, though

Table 1. Experiment One – Mean time to complete task (seconds)

	Younger faces	Older faces	Overall
Younger adults	16.83	30.26	23.54
Older adults	24.40	30.00	27.20
Overall	20.61	30.13	25.37

Table 2. Experiment One – Analysis of variance summary table

A = the between-subjects variable (rows) older versus younger adults					
B = the repeated-measures variable (columns) older versus younger faces					
Source	SS	df	MS	F	P
A	468.11	1	468.11	1.6	0.210
B	3168.26	1	3168.26	16.74	0.0001
A x B	536.26	1	536.26	2.83	0.097

this did not reach conventional statistical significance (F = 2.83, df 1, 68, p = 0.097, n.s.). For subsequent experiments, an inspection of the above mean scores suggests that a more powerful experimental design might detect that the younger adults were faster than older adults for younger faces but not at all faster for older faces.

Errors were recorded, but were few occurred and no significant differences were seen. This indicates that the participants were acceptably successful in following the instructions to work carefully. Finally, analyses of the learning curves were predominantly power functions for both the younger adults (F = 15.71, d.f. 1,8, p < 0.005) and for the older adults (F = 13.86, d.f. 1,8, p < 0.01). These curve fitting analyses will be reported in more depth in the longer version of this paper.

2.6 Discussion

This initial experiment provides a number of very important leads for our second experiment. First, our participants showed substantial learning, in the form of a power function that indicates moderate levels of motivation, as shown in other, ongoing work in the CIRCUA lab. (A report of this work is available on request). Second, the older faces took more time to process than younger faces, though we have confounded type of face with practice level, as this gives a simpler design to allow a clearer view of the resulting learning curves. On the basis of the present experiment, therefore, it is possible that the observed types-of-face difference is due, at least in part, to practice. This point will be examined in the second experiment. Third, our participant made few errors, suggesting that they were reasonably well motivated and that speed–error trade-offs do not significantly contaminate the observed performance times. Fourth, perhaps most intriguing, the data suggest, but do not prove, an own-age bias. In these data, younger adults were better than older adults than with younger faces, but for older faces younger and older adults performed at the same level, with a hint that older adults could perform better than younger adults. The best response to such speculations is a further experiment and that is the next step.

3 Experiment Two

3.1 Introduction

The second experiment builds significantly on the evidence provided by the first study. The present work again involves the creation of two types of pass-face prototype systems, one using younger faces and one using older faces, used by a sample of older adults (over 60 years of age) and younger adults (under 30 years of age). In the second experiment, order of testing of the different conditions was counter-balanced across participants. The quantity of pre-test practice was increased to four trials, as there as some indication that the participants in experiment one were slower on their first trial, as they were still inclined to study the instructions one more time. Finally, the size of the pass-face key (equivalent to password in a conventional system) was varied, such that half the pass-face keys were made up of two faces and the other half were based on four faces.

3.2 Participants

The participants were in two samples, (a) older adults (over 60 years of age) and (b) younger adults (under 30 years of age). There were eight participants in each group. The age of each person was recorded but their identities and other demographic details will not (e.g. occupation, location etc) in order to protect their anonymity. Participants were recruited from local sources, subject to the following conditions. No copyright is involved, no famous people are involved and no people judged to be unique, very unusual or very easy to recognise were included. Each participant worked alone.

3.3 Materials

The pass-face systems were based on arrays of nine faces at a time, on the screen of a standard laptop, based on simple word processed documents, with included hyperlinks for target words and faces. Half the arrays used younger faces (under 30 years) and the other half were based on older faces (over 60 years of age). Age groups were not mixed within arrays. For each prototype, a pass-face key (a pass-face equivalent to a password) was created. To form the pass-face keys, two or four faces were chosen at random from the sample of younger faces and two or four faces were chosen at random from the sample of younger faces. The pass-face sequence alone was presented before each prototype. Subsequently, a random array of nine faces was presented on one screen, within which the pass-face items were included at random locations. The background items were selected at random from the sample of faces of the correct age group, excluding the current pass-face sequence. Next, a second array of nine faces was presented on one screen, within which the pass-face items were included at random locations. Each face was linked by a hyperlink to an appropriate response page that provided feedback to the participant.

3.4 Methods

At the start, each participant was given four practice sessions in which an incorrect response took them back to the initial pass-face sequence on a new screen that contained the message "Apologies, you will have to start again." Order of testing was counterbalanced across participants as shown in table three. The instructions emphasised accurate performance, asking participants to work as quickly as possible without making errors.

Each participant was tested alone. They were each shown a pass-face key of two faces for five minutes and asked to study the faces so that they would be able to recognise them in a bigger set of faces (nine at a time). They were then shown a display of nine faces that included the pass-face sequence items at randomly chosen locations. The participant then selected any one of the two pass-face key items by cursor clicks. As each face depiction was a hyperlink, the act of selection of a face took the participant to a new screen. If their response had been correct, the new screen would contain the message "Well done, that was correct. Now select another one of your faces" and an array of nine faces would be displayed. This process continued until the participant correctly selected both pass-face key items. At this point, the participant was able to a simple Internet based selection task where they could select screens that reflected their hobbies and interests. If the wrong face was selected, the participant was redirected to a different screen that simply said "Apologies, you will have to start again."

All participants were told that the purpose of the work was to evaluate the two prototypes (Older faces or younger faces) to see which, if any were better. It was emphasised that the prototypes were under test and not the participants themselves. All participants were told that they would try out the prototypes a number of times but the exact number of times, so as to avoid end-of-session effects. All participants were told about the idea of the pass-face concept and to expect either two faces or four faces. They were also told that some collections of faces would be older faces and other collections were of younger faces. They would always select from nine faces and those sets faces were always older or younger faces but never mixed age. Samples of older faces (over 60) and younger faces (under 30) were collected from local sources, subject to the following conditions. No copyright is involved, no famous people are involved and no people judged to be unique, very unusual or very easy to recognise were included.

Since the study involved two types of pass-face; one was based on two faces and one based on four faces. Half the trials used two faces and half the trials used four faces. The pass-face sequences were used in two blocks of trials, such that the first half of the session used one type of pass-face sequence, whilst the second half used the other type of pass-face. Half the participants received the two face version first and half received the four face version first. The trials were also divided into older faces and younger faces. Within the two blocks, half the participants received the same age-group faces first and the other half received the different age-group first. The design is shown by the following table. Note that this design allows for analysis in terms of both "same-age versus different age" and "older faces versus younger faces" Each block of trials consisted of four trials, each with a new pass-face sequence each time.

Measures of performance were; time to complete a trial and number and types of errors made.

Table 3. Experiment three counter-balancing conditions and order of testing

Group	Practice block	First block	Second block	Third block	Fourth block
1	Four practice sessions	2 faces different age	2 faces different age	4 faces same age	4 faces different age
2	Four practice sessions	2 faces different age	2 faces same age	4 faces different age	4 faces same age
3	Four practice sessions	4 faces same age	4 faces different age	2 faces same age	2 faces different age
4	Four practice sessions	4 faces different age	4 faces same age	2 faces different age	2 faces same age

Table 4. Experiment two; average times in seconds

YOUNG FACES	YOUNG FACES	OLD FACES	OLD FACES	
BIG GRIDS	SMALL GRIDS	BIG GRIDS	SMALL GRIDS	
113.13	131.13	119.00	100.88	YOUNGER ADULTS
212.25	135.75	166.13	173.75	OLDER ADULTS

3.5 Results

Inspecting the above data, it is clear that for big grids (four faces), older adults are faster with older faces than with younger faces. Also, for big grids, younger adults are faster with younger faces than with older faces. Younger people are faster than older people for both big grids and small grids (two faces). Turning to small grids, older adults are faster with younger faces and younger adults are faster with older faces. This pattern of results indicates a three-way interaction between age of participant, age of faces and size of pass-face key. These data were analysed by means of a three way analysis of variance (ANOVA), with type of participant (between subjects), type of pass-face (within subjects) and size of pass-face size (within subjects). Here is the summary of the ANOVA:

3.6 Discussion

These are important results for the concept of universal accessibility. Whilst younger adults performed better overall than older adults in experiment two, this advantage was lost in experiment one with older faces. The choice of older versus younger faces made a very significant difference in both studies, so should not be ignored when designing pass-face systems for accessibility and security. However, the three significant interactions complicate the picture in experiment two. The age of participants and the age of the pass-faces interacted with each other, as did grid size and age of faces. However, the most striking interaction was the three way interaction between the three factors. This demonstrates that these three factors cannot be ignored when designing pass-face systems, unless you limit your intended users to one age group. It also means that whilst one design may be more accessible for younger users (young faces), a completely difference design would be more accessible for older users (older faces). On that basis, we can conclude that (a) older adults and younger adults have quantitatively different recognition memories and (b) the significance of these results means that the definition of accessibility reflects both the type of user and the task context.

Table 5. Experiment two; analysis of variance (significant factors in bold)

factor	type	description	F value	df	significance level (two tailed)
A	main effect	Age of participants (between subjects)	4.81	1,14	$p < .05$
B	main effect	Grid size (within subjects)	1.96	1,14	n.s.
AB	interaction	Participant age & grid size	1.96	1,14	n.s
C	main effect	Age of pass-faces (within subjects)	3887.51	1,14	$p < 0.001$
AC	interaction	Age of participants & age of faces	3884.07	1,14	$p < 0.001$
BC	interaction	Grid size & age of faces	50.59	1,14	$p < 0.001$
ABC	interaction	Age of participants & grid size & age of faces	50.96	1,14	$p < 0.001$

References

1. Bahr, G.S., Ford, R.: Why Pop-Ups Don't Work And How To Make Them Effective: Pop-Up Prompted Eye Movements, User Affect And Decision Making. Computers in Human Behavior 27, 776–783 (2011)
2. Barrington, L., Yoder-Wise, P.: Executive control function: a clinically practical assessment. Journal of Gerontological Nursing 32(2), 28–34 (2006)
3. Elias, J.W., Treland, J.: Executive function in Parkinson's disease and subcortical disorders. Seminars in Clinical Neuropsychiatry 4(1), 34–40 (1999)
4. Levitt, T., Fugelsang, J., Crossley, M.: Processing speed, attentional capacity, and age-related memory change. Experimental Aging Research 32, 263–295 (2006)
5. Park, D.: Consumer Fraud and the Aging Mind. Scientific Testimony Presented to the Senate Special Committee on Aging (July 27, 2005), `http://www.centerforhealthyminds.org/downloads/park%20testimony.pdf` (accessed July 2007)
6. Perfect, T.J., Moon, H.C.: The own-age effect in face recognition. In: Duncan, J., Phillips, L., McLeod, P. (eds.) Measuring the Mind, pp. 317–337. Oxford University Press, Oxford (2005)
7. Royall, D.R.: Précis of executive dyscontrol as a cause of problem behavior in dementia. Experimental Aging Research 20, 73–94 (1994)
8. Skurnik, I., Yoon, C., Park, D.C., Schwarz, N.: How warnings about false claims become recommendations: Paradoxical effects of warnings on beliefs of older consumers. Journal of Consumer Research 31, 713–724 (2005)
9. Stableford, S., Mettger, W.: Plain Language: A Strategic Response to the Health Literacy Challenge. Journal of Public Health Policy 28, 71–93 (2007)
10. U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. Quick Guide to Health Literacy, `http://www.health.gov/communication/literacy/quickguide` (accessed July 2007)