

Eye Tracking-based Evaluation of User Engagement with Standard and Personalised Digital Education for Diabetic Patients

Susan Quinn
Ulster University
Jordanstown, BT37 0QB
quinn-s47@ulster.ac.uk

Raymond Bond
Ulster University
Jordanstown, BT37 0QB
rb.bond@ulster.ac.uk

Chris Nugent
Ulster University
Jordanstown, BT37 0QB
cd.nugent@ulster.ac.uk

Health education is a fundamental process in strategies to inform and involve patients with effective health management. Nevertheless, challenges exist in motivating patients to engage with health information. This paper describes a study which evaluated user engagement with generic and tailored digital formats of interactive health education for diabetic patients. We recruited 18 participants, nine of whom viewed a diabetic education booklet, and the remaining subjects viewed electronic personalised education. Eye tracking was used to measure visual engagement with the artefacts, and the users provided feedback regarding the aesthetics and usability of the education. The findings indicated that for both artefacts the participants attended predominantly to the text content. Moreover, for the personalised education, it was found that images that had been personalised in accordance with user characteristics received a greater amount of attention than non-tailored images.

Personalisation, Patient Education, Engagement, Eye Tracking

1. INTRODUCTION

Having access to health information can assist an individual to become more knowledgeable about their health needs and better able to exercise informed choices about their healthcare [1, 2]. However, whilst access to printed and online health education has proliferated, many health consumers experience reduced ability to engage with the information acquired [1]. *Health literacy* is a set of communication and cognitive skills which assist an individual to use health information to make daily and longer-term health decisions [3, 4]. Nevertheless, while health literacy may impact an individual's capability to use health information, the quality attributes of health education artefacts may also reduce their comprehensibility for consumers [5]. For many patients printed educational materials such as leaflets and booklets provide a trusted form of health information [5, 6]. However, studies have found that the content, format and readability of written artefacts do not always match the information needs and literacy competencies of all patients [7, 8]. There are concerns that if educational materials are incongruent with patients' literacy skills then some patients may misunderstand the information given or disengage with further healthcare [9].

In order for a patient to process and subsequently recall health information they must be attentive to the information provided [10, 11]. Nevertheless, a challenge exists in identifying methods of providing

health information that will encourage patients to engage with the material. Reviews of health behaviour change interventions have described how tailored health communications have been effective in motivating health consumers to engage with positive behavioural changes [12, 13]. One means by which tailoring could be applied to attract the attention of a patient is to present health education as self-referential. *Personalisation* is a technique that incorporates recognisable characteristics of an individual (e.g. an individual's name) into a health message in order to enhance the relevance of the information for that person [11, 14]. Perceiving a message as self-relevant may intensify central processing of the message by an individual [11, 14].

This paper describes a study which explored the engagement behaviour exhibited by a user when viewing either a generic education booklet for diabetic patients or web-based personalised diabetic patient education. Diabetes is a chronic condition in which the body may not produce sufficient amounts of the hormone insulin, or may not be able to use the produced insulin effectively [15, 16]. The medical complications of diabetes can include cardiovascular disease and neuropathy, moreover the condition can have a psychological impact on an individual resulting in increased stress or depression [15, 16]. The prevalence of diabetes is increasing worldwide as it is estimated that 415 million people are currently diagnosed as diabetic [15]. Managing diabetes can be a demanding

responsibility for the patient and one fundamental approach to assisting patients is through the provision of high-quality education [15, 17]. During this study, groups of participants read either a booklet that was published by a registered UK charity, *Diabetes UK* [16], or web-based personalised education as provided by the study authors. A main objective of the study was to identify whether there were any similarities or contrasts in the engagement patterns of individuals when using the different artefacts. Viewing behaviour was recorded and utilised, along with the participants' self-measures of engagement, to uncover patterns of engagement with the two formats of education. The remainder of the paper is structured as follows. Section 2 discusses related research in which eye tracking has been used to examine how health consumers read and utilise health information. Section 3 describes the methodology used in the study. The results of the study are reported in Section 4 and further discussed in Section 5.

2. RELATED WORK

User Engagement is a multifaceted concept for which various definitions, activities and behaviours have been proposed [18]. For example, it has been suggested that user engagement is revealed through a user's interest and attention towards technology [18]. O'Brien and Toms [18, 19] proposed that engagement "*is a quality of user experiences with technology*" which is characterised by various dimensions including aesthetic and sensory appeal, interest, motivation and usability. They developed the *User Engagement Scale (UES)* [20], a self-report instrument by which users could indicate their engagement with a system with relation to aesthetic appeal, felt involvement, perceived usability, focused attention, endurability and novelty. The UES has been validated as a reliable measure of engagement in various online platforms including news and ecommerce [20, 21].

A contrasting approach to evaluating an individual's attention towards health information is through analysis of eye gaze activity. Eye tracking is a technique for documenting and quantifying eye movements [22], and has been employed to appraise users' focus of attention and cognitive load when searching for and using health information [10, 23, 24]. Eye movement data can indicate where the user's interest is directed on a visual display and how the level of interest fluctuates in response to different visual stimuli. Different types of eye movements are evaluated to identify the pattern of user gaze behaviour. The most prominent types of movement used include, (1) a *fixation*, during which the eye is stationary and focused on a specific area of a visual display, and (2) a *saccade*, in which the eye moves rapidly between fixations. An individual will perceive visual images only during periods of

fixations, and are wholly unable to see during saccades. It has been proposed that the length of a fixation is indicative of the level of cognitive processing being applied towards an object [22, 25].

Health risk information is often displayed on the packaging of consumer goods such as alcohol. Consequently, much research has focused on examining whether this information is noticeable, and is read by consumers. An eye tracking investigation of attention towards brand information and health information on alcohol and soda packaging found that the subjects spent a larger amount of time viewing branding information [23]. Nevertheless, for all the subjects the size and appearance of alcohol labels did have an impact on viewing time as warning labels that were larger and less complex were attended to for longer periods.

Eye tracking methodologies have also been utilised to investigate the associations between health literacy and the manner in which health information is used and recalled. Mackert et al. [24] used eye tracking to investigate whether there was an association between the information viewed on the Newest Vital Sign (NVS) nutritional label and health literacy competency. They found that subjects with reduced health literacy skills tended to fixate on non-relevant information for longer periods of time. A study which examined attention towards a text only webpage and a text-illustrated webpage found no association between health literacy competency and the fixation rate or fixation length for the two formats of health information [10]. However, it was found that increased attention to the text of the webpage correlated with increased recall of the information for those with adequate health literacy, but that an equivalent relationship was not found for those with limited health literacy. Conversely increased attention to the illustrations was associated with improved information recall for those with limited health literacy but not for participants that had adequate health literacy. The authors of [26] used eye tracking to examine the gaze behaviour of low literacy adults when viewing a Centers for Disease Control and Prevention (CDC) factsheet, or a literacy appropriate information decision aid. They found that the more densely worded factsheet was more challenging for some participants. Fixation analysis indicated that some participants tended to dwell on certain words and phrases for longer durations, whilst gaze patterns showed that some participants had an unsystematic approach to reading the text, repeatedly re-reading lines of text or looking at the text in a haphazard manner. A literature review did not identify any eye tracking studies that focused on diabetic patient education.

Health education comprises a process which may assist patients to become more activated in decisions about their health. The reviewed literature suggests that health consumers will employ different

methods to engage with health information. Moreover, some individuals will not be attentive to all the information provided. Therefore strategies are required that will motivate patients to attend to health education. This paper describes a study which appraised engagement with generic and personalised formats of diabetic patient education. User engagement is driven by various characteristics thereby a dual approach to measuring engagement was employed. The UES provided a means to gather subjective ratings of engagement from the participants themselves, whilst eye tracking afforded evaluation of the participants' visual responses to different aspects of the education. This enabled evaluation of the participants' experience of the visual appeal of the education, whilst also appraising which features of the education attracted most attention.

3. METHODOLOGY

During the study each participant attended an eye tracking experiment. In the first stage a participant completed demographic information including age, gender and highest qualification, and rated their knowledge of diabetes on a ten-point Likert scale of one to ten (ranging from very limited to expert). The participant also rated their information technology expertise on a ten-point Likert scale of one to ten (ranging from novice to expert). Following this the eye tracking began, and the participant read diabetic patient education (generic or personalised) which was presented on a computer screen. No time limit was set to complete reading the education. When reading was complete the participant was presented with four health questions, and was free to use the education to find answers for the questions. Following this the participant completed the UES.

3.1 Participants

A total of 18 subjects participated in the study, which comprised 10 (56%) male and eight (44%) female participants. The background disciplines of the participants included Computer Science (50%, $n = 9$), Politics (11.1%, $n = 2$), Museum Studies (5.6%, $n = 1$), Arts (11.1%, $n = 2$), English (5.6%, $n = 1$), and Cookery (5.6%, $n = 1$).

This was a between-subjects study in which a *control group* of nine participants viewed a *generic education booklet* for diabetic patients, and an *experimental group* of nine participants viewed *web-based personalised diabetic education*. Before the experiments began each participant was randomly assigned to one of these groups. The control group included six (66.7%) male and three (33.3%) female participants, with an age range of 24 to 36, and a mean age of 31.1 years ($SD = 4.31$). All of the members of this group had achieved an undergraduate or postgraduate degree. The

experimental group included four (44.4%) male and five (55.6%) female participants, with an age range of 19 to 49, and a mean age of 32.9 years ($SD = 9.61$). Six (66.7%) of the participants in this group had attained an undergraduate or postgraduate degree, and the remaining three participants (33.3%) had a Regulated Qualifications Framework qualification at level 2 or level 3 [27]. The range of information technology and diabetes expertise within the two groups was similar. The average diabetes knowledge score for the control group (3.33 out of 10) was greater than the average score for the experimental group (2.89 out of 10), however the difference between the two scores was not statistically significant ($t_{(16)} = -.476$, $p = .641$). Similarly, the mean score for information technology expertise was greater for the control group (6.56 out of 10) when compared to the experimental group (5.89 out of 10), however again the scores did not differ significantly ($t_{(16)} = -.514$, $p = .616$).

3.2 Diabetic Patient Education

3.2.1. Generic Patient Education Booklet

During the study the members of the control group viewed an introductory booklet for diabetic adults, *Understanding diabetes. Your essential guide* [28], which was published by *Diabetes UK*. This booklet is aimed at patients that have been diagnosed with type 1 or type 2 diabetes, or caregivers of diabetic patients. The participants were required to read a portion of the booklet which covered information related to the diagnosis, management and complications of diabetes. The pages read by the participants contained paragraphs of text and nine images. The content of these pages did not alter during the study thereby all the members of the control group were presented with the same text and images. The average number of words on a page was 225.22 words. The Simple Measure of Gobbledegook (SMOG) readability formula [29] was used to determine the readability of the text. A text with a higher SMOG score may be more challenging to understand for some individuals. The SMOG score was 9.8, and this readability level was understandable for all the members of the group.

3.2.2. Web-Based Personalised Patient Education

Members of the experimental group viewed web-based personalised education. The education was provided through an ontology-based architecture which was previously developed by the authors, and is described in [30]. The education consisted of five web pages which focused on the topics that were included in the generic booklet. The personalised education incorporated both *static* features which remained constant for all the participants, and *personalised* features, which varied between participants. For example, the layout of the web pages and the text, headers and sub-headers remained largely consistent for all the participants.



Figure 1: Screen shot of education web page that focuses on physical activity, and contains static, personalised and contextualised content.

Four of the web pages included at least one *static image* which remained unchanged for all the participants. The personalised features of the education were as follows.

- (i) The first line of the first web page comprised a personalised greeting which included the participant's name.
- (ii) Each of the five web pages contained at least one *personalised image* which was selected based on the participant's personal characteristics and preferences. For example, Figure 1 illustrates a web page which contains physical activity focused information for a female participant. The two personalised images are gender and age specific and are also contextualised to the preferred sporting activity of the participant.
- (iii) The *physical activity* web page contained information that had been personalised to the preferred sporting activity of the participant, and included a Google map that indicated a route from the participant's geolocation to a location for the activity.

In summary, the personalised education contained six static images and seven personalised images. The average number of words per web page was 356.6 words. The SMOG readability score for the education was 7.6, thereby the text content was comprehensible for all the members of the group.

3.2.3. Comparison of the Artefacts

The two artefacts had several comparable features. The content of each focused on the same diabetic health topics and both contained generic images that remained consistent for all participants. However, the SMOG scores indicated that the booklet had lower readability than the personalised education. Features that were unique to the personalised education included using the

participant's name, and the inclusion of images that reflected the gender, age group and preferences of the participant.

3.3 Eye Tracking

During an experiment a *Tobii X60 Eye Tracker* was used to record the eye movements of a participant. The eye tracker uses near infrared illumination to detect the user's corneal reflection and in turn identify gaze points on the screen, at time intervals specified by the sampling data rate (60Hz) [22, 31]. The collected data was transmitted to the *Tobii Studio* software application, where the *Tobii I-VT filter* [32] was applied to calculate various metrics.

3.4 User Engagement Scale

In the last stage of the experiment the participant completed the *User Engagement Scale (UES)* to report their engagement with the education. Related research had identified an association between perceived attractiveness of health information and an increased likeliness of attending to the information [33], thereby it was decided to focus on two dimensions of the participants' experience of the education; aesthetic appeal and usability. Two of the UES subscales, *Aesthetics (AE)* and *Perceived Usability (PUs)* were adapted for use in the study.

The AE scale focused on a participant's evaluation of the visual appearance of the education [19]. The scale comprised five items that addressed the following aspects of the education.

- (i) The aesthetic appeal and attractiveness of the education.
- (ii) Whether the education was visually pleasing and appealed to the participant's visual senses.
- (iii) Whether the participant liked the graphics and images included.

Each item was scored on a five-point Likert scale of one to five (ranging from strongly disagree to strongly agree) with a final sum score within the range of five to 25. A higher end score indicated that the education was perceived as highly aesthetically appealing.

The PUs scale addressed a participant's emotional and cognitive responses to the education [19]. The original scale contained seven items, however it was decided that only six of the items were applicable to the patient education. Thereby the scale used in the study consisted of six items that were scored on a five-point Likert scale of one to five (ranging from strongly agree to strongly disagree) with a final sum score within the range of six to 30. A higher end score indicated that the education was perceived as highly usable. The six items focused on the following qualities of the education.

- (i) Whether the education was confusing to use.
- (ii) Whether the participant felt frustrated, annoyed or discouraged while searching for information in the education.
- (iii) Whether the participant judged that searching for information in the education was mentally taxing or demanding.

3.5 Health Questions

After a participant had completed reading the education they were presented with four health questions, as listed in Table 1. The information required to answer the questions was available in both education artefacts. Presenting a participant with health questions obliged them to search through the education. This ensured that the participant would have experience of seeking information within the artefact and could effectively evaluate the perceived usability of the education as measured by the items of the PUs scale (as described in Section 3.4).

Table 1: Health questions presented to a participant during an eye tracking experiment.

	Health Questions
1	Which test is commonly used to indicate a patient's blood glucose levels over the past 2-3 months?
2	"One common symptom of diabetes is unexplained weight loss." Is this statement true or false?
3	What is the recommended minimum amount of moderate intensity exercise that adults should complete each week?
4	Hypoglycaemia can occur if a patient's blood glucose level falls below 4mmol/l. Can you suggest 1 reason why hypoglycaemia may occur?

3.6 Statistical Analysis

The eye tracking metrics for each group were compared using a *t*-test, or a Mann-Whitney U test for data that was not normally distributed. Correlations between various eye tracking measures and the UES engagement scores were investigated through use of Pearson's product-moment correlation, Spearman's correlation coefficient or Kendall's Tau.

4. RESULTS

All of the participants read the required portion of the education, and completed the four health questions and the UES. The mean experiment completion time for the experimental group was 738.56 seconds (SD 237.68), and the mean for the control group was 815.11 seconds (SD 250.47). A comparison of the mean completion times, using a *t*-test, did not find a significant difference between the two groups ($t_{(16)} = -.665, p = .515$). There was variation in the question scores achieved by the members of the experimental group; one participant (11.1%) achieved a score of two, while three participants (33.3%) achieved a score of three, and the remaining five participants (55.6%) answered all four health questions correctly. All of the members of the control group achieved a full score of four. Evaluation of the relationships between the health question scores and gaze behaviour found only one statistically significant relationship. For the experimental group a positive correlation was identified between question score and the total number of fixations on the text content ($r = .73, p < .05$). No relationships were identified between the question scores and the diabetes knowledge scores.

4.1 Engagement with the Generic Booklet and the Personalised Education

4.1.1. User Engagement Scale

The distribution of the aesthetics (AE scale) and perceived usability (PUs scale) scores for both formats of education are illustrated in Figure 2. For each participant, the five individual item scores were summed to provide an overall aesthetics score (out of 25). The scores for the personalised education ranged from 17 to 25, with a mean score of 21.2 (SD 3.31), and for the generic booklet the scores extended from 16 to 25, with a mean score of 20.56 (SD 3.08). Furthermore, *t*-test analysis did not identify a significant difference between the two sets of scores ($t_{(16)} = .442, p = .664$). The AE scale comprised five individual items which related to various attributes of engagement. The average scores for each of these items were calculated. For three of the items, aesthetic appeal, attractiveness and visually pleasing, the average score for the web-based personalised education was higher than the average for the generic booklet. However, for the

items related to the graphics and images, and the appeal of the education the average scores were equivalent.

For each participant, the six individual item scores were summed to provide an overall perceived usability score (out of 30). For the personalised education the perceived usability scores ranged from 22 to 30, with a mean score of 27.56 (SD 3.0), and for the generic booklet the mean score was 25.0 (SD 3.32) and the range was from 20 to 30. The two sets of scores were not significantly different ($U = 23.50, p = .136$). The average scores for the six PUs scale items were calculated. The results indicated that for all six items (confused, annoyed, frustrated, discouraged, mentally taxing, demanding) the generic booklet had a reduced score when compared to the corresponding score for the personalised education.

Kendall's Tau was used to examine the relationship between the aesthetics and perceived usability scores. The results for both the generic booklet ($\tau = .030, p = .914$) and the personalised education ($\tau = .459, p = .114$) indicated that no statistically significant relationships existed between the scores for the two factors of engagement.

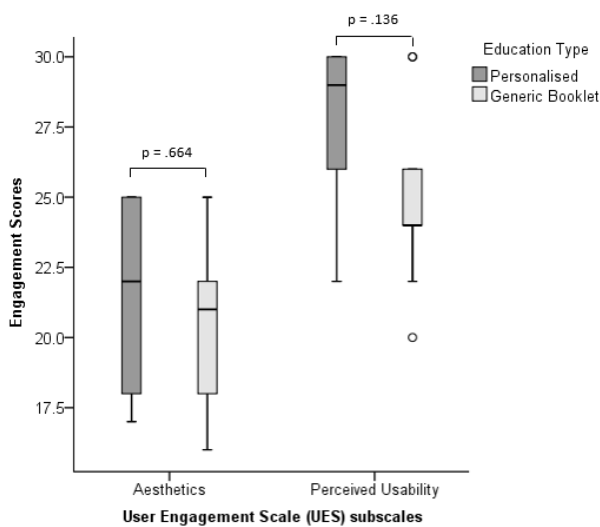


Figure 2: Boxplots of the scores for aesthetics (AE scale) and perceived usability (PUs scale) for the generic booklet and web-based personalised education.

4.1.2. Gaze Behaviour

The generic booklet and personalised education comprised two types of educational content which were text and images. However, while the images in the generic booklet would not alter throughout the experiment, for the personalised education the images were further categorised as, (1) *static images* which remained constant for all participants, and (2) *personalised images* that would change in accordance with the characteristics of a participant. The first stage of analysis focused on comparing the

overall gaze behaviour of the members of the two groups. *Tobii Studio* was used to define the text and image elements within the education as *Areas of Interest (AOI)*. Various fixation based metrics were then calculated for the AOI.

Figure 3 illustrates the proportion of fixations for the different types of content in each education artefact. For both artefacts the majority of fixations were directed at the text content. The percentage of fixations on image components was larger for the personalised education (6%) when compared to the booklet (2%), however this is not surprising as the personalised education contained a larger quantity of images. The booklet contained a larger total of static images ($n = 9$) than the personalised education ($n = 6$), however the proportion of fixations on the static images was equivalent (2%). While the personalised education contained a comparable amount of personalised ($n = 7$) and static images ($n = 6$), the percentage of fixations on the personalised images (4%) was double the proportion for the static images (2%).

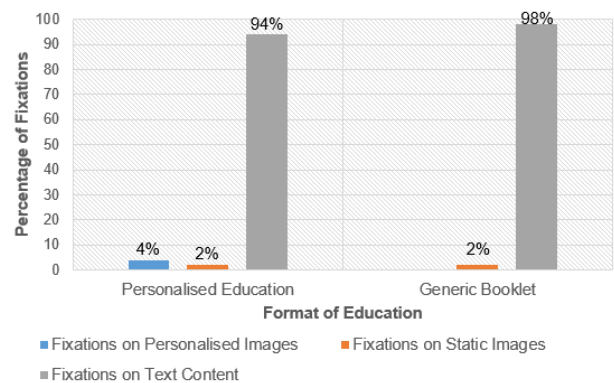


Figure 3: Bar chart indicating percentage of total fixations on different types of content for the generic booklet and personalised education.

The assessment of the overall gaze behaviour suggested that there may be a similarity between the two study groups in the overall proportion of attention directed towards static images. However, in order to clarify whether any similarities or differences existed in the engagement patterns it was decided to evaluate the fixation behaviour of the two groups in relation to, (1) the static images, and (2) the text content of the educational artefacts.

4.1.3. Static Images

Figure 4 indicates the distribution of the *total number of fixations* on the static images for the members of both groups. With an average of 32.56 fixations per participant (SD 16.76) the mean fixation rate for the control group was larger than the mean rate for the experimental group (19.0 fixations per participant, SD 12.67). T-test analysis did not find a statistically significant difference between the mean fixation

count for the two groups, however the low p -value suggests a trend and perhaps greater sample size and statistical power would yield significance ($t_{(16)} = -1.94, p = .071$).

It has been suggested that if elements of visual stimuli appear interesting, appealing, confusing or complex it may be more challenging for an individual to process these, and this may result in a longer period of fixation [22]. Evaluation of the fixation rate had suggested that the control group had attended more to the static images than the experimental group. However, the next stage of analysis attempted to quantify attention by using a temporal measure, *fixation duration*, which was defined as the length of a fixation in seconds. The average fixation duration per participant was larger for the control group (6.61 seconds, SD 4.18), when compared to the participants in the experimental group (4.1 seconds per participant, SD 3.53). However the mean fixation duration did not differ significantly between the two groups ($t_{(16)} = -1.38, p = .186$).

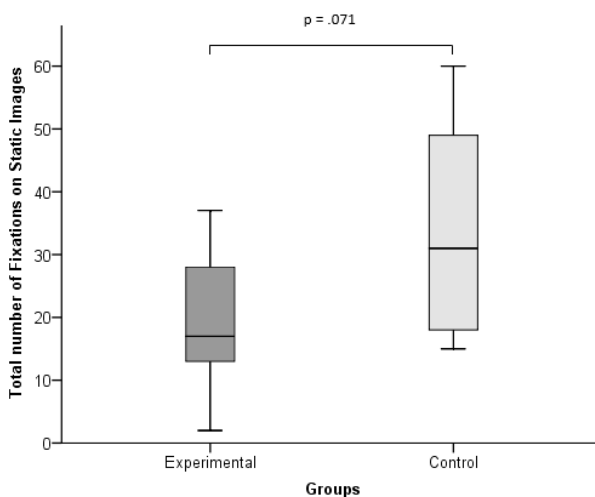


Figure 4: Boxplots indicating total number of fixations on static images for experimental and control group.

The metrics analysed to this point assess fixation on the static images. However, it is also useful to comprehend a participant's overall gaze behaviour on an AOI as this can reflect whether the participant returned to an AOI more than once while viewing the education. An AOI *visit* refers to a participant's observation of an AOI and includes both fixations and saccades. In Tobii Studio a *visit* to an AOI corresponds to the duration between the beginning of the first fixation on an AOI and the beginning of the first fixation outside that AOI. Evaluation of the visit data suggested that the participants in both groups did revisit the static images whilst reading the education. The mean number of visits to static images for a member of the experimental group was 12.67 visits (SD = 9.54), while for the control group the mean was much higher at 19.44 visits per

participant (SD 10.2). However there was no significant difference between the average visits for the two groups ($t_{(16)} = -1.46, p = .165$).

4.1.3.1. Relationships between engagement quantifiers

The eye tracking data signified the participants' level of interest in the static images, however it was decided to investigate how this data correlated with the participants' own appraisal of their engagement with the education (as denoted by the AE and PUs scores). Regarding the experimental group, no significant associations were found between the aesthetics scores and the perceived usability scores and any of the eye tracking metrics discussed. In contrast, for the control group statistically significant moderately positive relationships were identified between the aesthetics scores and the total number of fixations ($r = .808, p < .01$), total number of visits ($r = .793, p < .05$), and the total fixation duration ($r = .779, p < .05$) for the static images. However, no significant relationships were found between the perceived usability scores and the eye tracking data for this group.

With relation to the demographic characteristics of the participants, no statistically significant correlations were identified between age and any of the eye tracking metrics or UES scores.

4.1.4. Text Content

The generic booklet had a greater volume of text, and correspondingly the mean time spent fixating on the text for members of the control group (352.1 seconds per participant, SD 183.86) was larger than the mean time interval for members of the experimental group (221.35 seconds per participant, SD 125.91). However, a t -test comparison of the data for the two groups did not find a statistically significant difference ($t_{(16)} = -1.76, p = .097$). Spearman's correlation was used to determine whether any relationships existed between diabetes expertise and total length of time spent fixating on the text. However, no statistically significant correlations were found either for members of the control group ($r_s = -.101, p = .795$) or the experimental group ($r_s = -.463, p = .209$).

4.2 Gaze Pattern Analysis for the Experimental Group

4.2.1. Personalised and Static Images

A key objective of the study was to examine whether the participants that viewed the personalised education were attentive to the personalised features of the education. The inclusion of images that were tailored to correspond with a participant's demographic characteristics comprised the most prominent visual personalised features. Thereby primary analysis focused on assessing whether there was a difference in the gaze behaviour for the personalised and static images. Figure 5 depicts the

distribution of the *total fixation rate* for the personalised and static images for the members of the experimental group. There was variation in the total number of fixations on the two types of images. The mean fixation rate for the personalised images (42.78 fixations per participant, SD 29.81) was more than double the mean rate for the static images (19.0 fixations per participant, SD 12.67). Furthermore, *t*-test analysis indicated that the fixation rate for the personalised images was significantly larger when compared with the rate for the static images ($t_{(8)} = 2.41, p < .05$).

The average fixation duration on the personalised images was 8.58 seconds per participant (SD 6.67), which was twice the length of the mean fixation duration on the static images (4.09 seconds per participant, SD 3.53). T-test analysis indicated that the difference between the duration means was very close to significance ($t_{(8)} = 2.29, p = .051$), however the result was not significant at the .05 level of significance. The mean number of *visits* per participant to the personalised images was 20.78 visits (SD 12.47), which was larger than the mean visit count for the static images (12.67 visits per participant, SD = 9.54). Moreover, this difference was statistically significant ($t_{(8)} = 2.36, p < .05$).

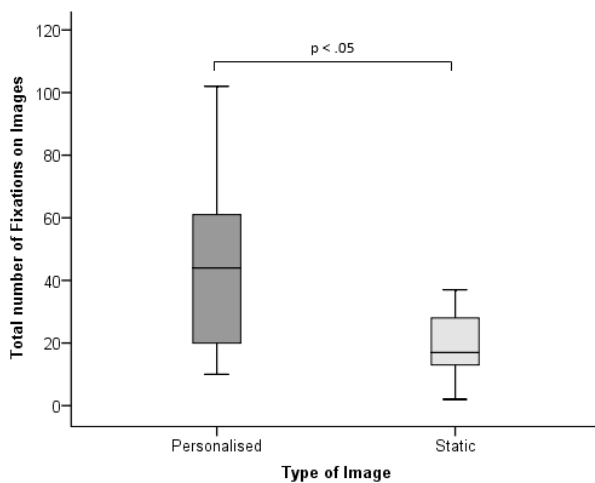


Figure 5: Boxplots indicating total fixations on personalised and static images for experimental group.

With relation to the UES scores, no statistically significant correlations were found between the scores for the aesthetics and perceived usability of the personalised education and any of the eye tracking measures (fixation count, fixation length, visit count).

5. DISCUSSION

Current approaches to patient education include the provision of standardised printed materials to all patients regardless of their demographics such as

age and cognitive abilities. However, whilst these materials provide medically verified advice, the information provided may not be accessible for all patients. This suggests that an adaptable approach to providing health education may be beneficial for patients and physicians. Strategies to enhance the efficacy of health communications have included the tailoring of electronic health interventions as appropriate for various demographic, behavioural and health attributes of consumers [34]. Personalisation is a tailoring strategy which attempts to increase the self-relevance of information for an individual, and has been proposed as a means to focus attention on, and encourage central processing of health messages [11, 12]. This paper described a study that evaluated the engagement patterns of users when viewing either a generic or personalised format of health education. Similarities and contrasts were identified in the engagement behaviours of the two groups of participants.

5.1 Engagement Behaviour of the Control Group and Experimental Group Members

The first phase of analysis focused on evaluating engagement with generic and personalised diabetic education. Two groups of participants viewed either a generic booklet or web-based personalised education, and engagement patterns for the two groups were compared. The subjects completed the UES, a self-reporting tool through which they rated their experience for two factors of engagement, the aesthetics and perceived usability of the education. An unobtrusive eye tracking device was also used to capture the gaze behaviour of the participants as they read the education. Thereby both subjective ratings and objectively gathered data was assessed. The aesthetics scores for the two formats of education were comparable, however examination of the individual items that comprised the aesthetics scale identified that for three items, aesthetics appeal, attractiveness and visually pleasing, the personalised education achieved a higher average rating. Despite differences between the two groups in the amount of time spent viewing the images, the average score for "liking" the images was equivalent. The average score for perceived usability was higher for the personalised education, however the scores did not differ significantly between the two types of education. The individual items that comprised the perceived usability scale related to the participants' experiences of searching through the education and included emotions such as frustration, annoyance, discouragement and cognitive demand. For all six scale items the generic booklet received less positive ratings when compared to the scores for the personalised education. This suggests that searching for information in the generic booklet was more challenging in comparison with the personalised education. Nevertheless, despite the variance in the

perceived usability scores all the participants in the control group correctly completed the four health questions, whilst only five members (55.6%) of the experimental group achieved a full score. It was unclear why there was a distinction in question scores between the two groups. The average diabetes knowledge score for the control group was greater than the mean score for the experimental group, and it is possible that previous knowledge may have contributed to the success of the experimental group. However, there was no significant difference between the mean scores for the two groups, thereby it is impossible to verify whether knowledge of diabetes had an impact on the health question scores for the control group.

Gaze behaviour metrics were used to identify where the visual attention of the participants was located on the education. For the members of both groups most attention was focused on the text content. A comparison of fixation on the text content suggested that the members of the control group spent more time reading the text when compared to the experimental group. However, the generic booklet contained a larger volume of text which may have contributed to the difference in fixation time. There are several reasons why both groups may have spent such a large proportion of their attention on the text. If visual information requires increased cognitive processing then an individual may spend an increased amount of time looking at the information [22, 25]. The ratings for diabetes knowledge indicated that 15 participants (83.3%) had rated their knowledge in the lower half of the scale, thereby it is likely that reading new information and encountering unfamiliar terminology may have resulted in increased attention. However, no relationships were found between level of diabetes expertise and amount fixated on the text. Moreover, it is also probable that the subjects spent longer reading the text content as they were aware that they would be answering questions about the information in the second stage of the experiment.

Appraisal of overall gaze behaviour indicated that the percentage of fixations on static images was equivalent for both groups. However, examination of the data indicated that the participants that viewed the generic booklet fixated on the static images more frequently and for longer time intervals when compared to the experimental group. Moreover, the members of the control group tended to revisit the static images more frequently than members of the experimental group. This behaviour may have been due to the types of images included in the booklet. The images in the generic booklet included a full-page diagram which explained the function of insulin. The eye tracking data suggests that this diagram was more successful than any other image in the booklet in capturing the attention of the participants (receiving 55% of all fixations). The diagram can be classified as a *cognitive illustration*,

an explanatory image which complements information and assists readers to more fully understand the text content [35, 36]. The remaining images in the booklet and all the static images in the personalised education were *affective illustrations*. These are illustrations that are used to draw attention to information, augment the attractiveness of the content and enhance the enjoyment of the information for the reader [35, 36]. Cognitive illustrations have been found to attract more attention than other illustrations in some online health information materials [35]. Similarly, in this study the diagram was especially successful in capturing the interest of the control group. Hence, it is probable that the increased level of interest for the diagram contributed to the increased attention towards the static images. Correlation evaluation highlighted an association between increased fixation on the static images by members of the control group and increased aesthetics scores for the generic booklet. However, a similar relationship was not identified for the subjects that viewed the personalised education.

5.2 Attention Towards the Web-Based Personalised Education

The second stage of analysis concentrated on the participants that viewed the personalised education. Viewing behaviour suggested that these participants looked at the personalised images more often and for double the length of time than was spent fixating on the static images. The personalised images had been matched to the participant's gender and age group, and there are several reasons why these images derived increased notice. A recent study of engagement with images on Instagram found that photos containing faces were more engaging, receiving more likes and comments, in comparison to photos not containing a face [37]. Thereby as most of the personalised image content represented people, while the static images focused on medical devices and food items, this may have encouraged engagement with the personalised images. Moreover, it is often found that humans have a preference for focusing their attention on people within pictures. For example, during an experiment which presented image pairs comprising a person-present and person-absent scene, the subjects fixated on the person-present scene first, and for a longer period of time [38]. Participants also have an increased likelihood to quickly locate and fixate on people when depicted in social scenes [39, 40]. Furthermore, several eye tracking studies have identified attentional bias towards human faces in pictures, with a particular focus on the eyes [40, 41]. It has been proposed that enhanced interest in the eyes is due to a tendency for humans to try to infer social information from an individual's eye gaze, and thereby better comprehend the meaning of a scene [39, 42]. It is evident that when viewing images,

humans tend to focus their attention on meaningful objects such as people. Thereby, as all of the personalised images contained people, this may have naturally attracted the attention of the participants. Consequently, it is probable that this contributed towards the increased focus towards the personalised images.

6. LIMITATIONS

The primary limitation of this study is related to the reduced size of the study groups. The volume of eye tracking data collected and the UES based data did provide an adequate amount of data to complete statistical analysis of various attributes of engagement. Moreover, statistically significant and non-significant relationships between engagement attributes were identified for the study population. However, an increased sample size may have highlighted further patterns in the gaze behaviour of the participants, or may have increased the statistical significance of the identified relationships.

7. CONCLUSION

The paper describes a study which investigated whether any relationships exist between the format of education being used and the engagement behaviour exhibited by an individual. The study evaluated characteristics of engagement for participants that used either a generic education booklet or web-based personalised education. A comparison of overall viewing behaviour determined some similarities between the two groups as it was established that most of the participants' attention was directed towards the text information and less attention was directed towards the images. Moreover, with regard to the participants' experience of the usability and visual appeal of the education no significant difference was found between the two formats of education. Statistical evaluation indicated that attention towards the static images was positively associated with the aesthetic appeal of the generic booklet. This would suggest that for some individuals the inclusion of images may contribute to the visual appeal of health education. However, for both forms of education no associations were found between the usability of the education and the amount of attention directed towards the text and images components.

Analysis of the gaze behaviour of the participants that viewed the personalised education indicated that the attention directed towards the personalised images was twice the amount than was directed towards the static images. Thereby, personalisation of the image content was successful in attracting the attention of the participants. Variations within the gaze behaviour revealed that the participants attended more to images which focused on faces.

The results of the study suggest that there are similarities and distinctions in the manner in which health consumers use different formats of health education. The viewing behaviour indicated that whichever format of education was used, the text content was the primary focus for the participants. This underlines that text content is the central element of a health education artefact for many users and draws most attention. Thereby one approach to improving the usefulness of health education is to ensure that the readability of the text information is comprehensible for the target audience, and focuses on their health information needs. For some of the participants an association was established between aesthetic engagement with the material and the inclusion of images. This indicates that for some individuals the inclusion of images in health education may help to motivate engagement with the material. The viewing behaviour suggests that the inclusion of image content that focuses on persons that a user can identify with may attract the attention of the user. The results indicate that image content can successfully attract the interest of a reader, and this provides an opportunity to improve the efficacy of patient education. Including images that reflect the characteristics and context of the user may encourage increased attention and assist users to identify and engage with the material. Moreover, utilising tailored images to complement the educational focus of the material may assist to inform the user, and enhance the efficacy of the education.

7.1 Future Research

The use of eye tracking and the UES tool provided different measures of user engagement. However, user engagement is a multifaceted concept which is challenging to quantify. Accordingly, future research could include an additional method for evaluating participant engagement such as introducing a device that measures physiological signals. For example, using a sensor device to measure heart rate variability or galvanic skin response would facilitate the measurement of natural physiological reactions to the education. This would augment the types of engagement behaviours recorded and enable a multifaceted evaluation of user engagement to be undertaken.

8. REFERENCES

- [1] Colledge, A., Car, J., Donnelly, A., and Majeed, A. (2008) Health information for patients: time to look beyond patient information leaflets. *J. R. Soc. Med.*, 101(9), pp. 447-453.
- [2] Speirs, S., Rees, S., and Tagoe, M. (2008) An audit of foot surgery information leaflets from the patients' perspective. *The Foot*, 18(1), pp. 7-14.

- [3] Sørensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska Z., and Brand, H. (2012) Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health*, 12, pp. 1-13.
- [4] Nutbeam, D. (2008) The evolving concept of health literacy. *Soc. Sci. Med.*, 67(12), pp. 2072-2078.
- [5] Shaw, A., Ibrahim, S., Reid, F., Ussher, M., and Rowlands, G. (2009) Patients' perspectives of the doctor-patient relationship and information giving across a range of literacy levels. *Patient Educ. Couns.*, 75(4), pp. 114-120.
- [6] Pohjanoksa-Mäntylä, M., Bell, J. S., Helakorpi, S., Närhi, U., Pelkonen A., and Airaksinen, M.S. (2011) Is the Internet replacing health professionals? A population survey on sources of medicines information among people with mental disorders. *Soc. Psychiatry Psychiatr. Epidemiol.*, 46(5), pp. 373-379.
- [7] Kondilis, B. K., Akrivos, P. D., Sardi, T., Soteriades, E., and Falagas, M. (2010) Readability levels of health pamphlets distributed in hospitals and health centres in Athens, Greece. *Public Health*, 124(10), pp. 547-552.
- [8] Haw, C., and Stubbs, J. (2011) Patient information leaflets for antidepressants: Are patients getting the information they need? *J. Affect. Disord.*, 128(1-2), pp. 165-170.
- [9] Svider, P. F., Agarwal, N., Choudhry, O. J., Hajart, A. F., Baredes, S., Liu, J. K., and Eloy, J. A. (2013) Readability assessment of online patient education materials from academic otolaryngology-head and neck surgery departments. *Am. J. Otolaryngol.*, 34(1), pp. 31-35.
- [10] Meppelink, C. S., and Bol, N. (2015) Exploring the role of health literacy on attention to and recall of text-illustrated health information: An eye-tracking study. *Comput. Hum. Behav.*, 48(7), pp. 87-93.
- [11] Hawkins, R. P., Kreuter, M., Resnicow, K., Fishbein, M., and Dijkstra, A. (2008) Understanding tailoring in communicating about health. *Health Educ. Res.*, 23(3), pp. 454-466.
- [12] Noar, S. M., Benac, C. N., and Harris, M. S. (2007) Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol. Bull.*, 133(4), pp. 673.
- [13] Krebs, P., Prochaska, J. O., and Rossi, J. S. (2010) A meta-analysis of computer-tailored interventions for health behavior change. *Prev. Med.*, 51(3-4), pp. 214-221.
- [14] Dijkstra, A. (2008) The psychology of tailoring-ingredients in computer-tailored persuasion. *Social and Personality Psychology Compass*, 2(2), pp. 765-784.
- [15] International Diabetes Federation. (2015) IDF Diabetes Atlas - Seventh Edition. <http://www.diabetesatlas.org/> (retrieved date: March 27, 2018).
- [16] Diabetes UK (2017) Know Diabetes. Fight Diabetes. <https://www.diabetes.org.uk/> (retrieved date: March 27, 2018).
- [17] Coppola, A., Sasso, L., Bagnasco, A., Giustina, A. and Gazzaruso, C. (2016) The role of patient education in the prevention and management of type 2 diabetes: an overview. *Endocrine*, 53(1), pp. 18-27.
- [18] O'Brien, H. L., and Toms, E. G. (2008) What is user engagement? A conceptual framework for defining user engagement with technology. *J. Am. Soc. Inf. Sci. Technol.*, 59(6), pp. 938-955.
- [19] O'Brien, H. L., and Toms, E. G. (2013) Examining the generalizability of the User Engagement Scale (UES) in exploratory search. *Information Processing & Management*, 49(5), pp. 1092-1107.
- [20] O'Brien, H., and Cairns, P. (2015) An empirical evaluation of the User Engagement Scale (UES) in online news environments. *Information Processing & Management*, 51(4), pp. 413-427.
- [21] Cano, M. B., Perry, P., Ashman R. and Waite, K. (2017) The influence of image interactivity upon user engagement when using mobile touch screens. *Comput. Hum. Behav.*, 77, pp. 406-412.
- [22] Tobii Technology AB. (2010) Tobii Eye Tracking. An introduction to eye tracking and Tobii Eye Trackers. <http://acuity-ets.com/downloads/Tobii%20Eye%20Tracking%20Introduction%20Whitepaper.pdf> (retrieved date: June 16, 2017).
- [23] Kersbergen, I., and Field, M. (2017) Alcohol consumers' attention to warning labels and brand information on alcohol packaging: Findings from cross-sectional and experimental studies. *BMC Public Health*, 17, pp. 123.
- [24] Mackert, M., Champlin, S. E., Pasch, K. E., and Weiss, B. D. (2013) Understanding Health Literacy Measurement Through Eye Tracking. *J. Health Commun.*, 18 (Suppl. 1), pp. 185-196.
- [25] Poole, A., and Ball, L. J. (2006) Eye tracking in human-computer interaction and usability research: current status and future prospects. In Ghaoui, C. (ed) *Encyclopedia of Human-Computer Interaction*. IGI Global.
- [26] Bass, S. B., Gordon, T. F., Gordon, R., and Parvanta, C. (2016) Using eye tracking and gaze pattern analysis to test a "dirty bomb" decision aid in a pilot RCT in urban adults with limited literacy. *BMC Med. Inform. Decis. Mak.*, 16, pp. 67.
- [27] Quality Assurance Agency for Higher Education, Office of Qualifications and Examinations Regulation, Scottish Credit and Qualifications Framework Partnership, Council for the

- Curriculum Examinations and Assessment Accreditation, Credit and Qualifications Framework for Wales, Quality and Qualifications Ireland. (2017) Qualifications can Cross Boundaries - A Guide to Comparing Qualifications in the UK and Ireland. Sixth Edition.
http://www.qaa.ac.uk/publications/information-and-guidance/publication?PubID=165#.WHbs_H3eJSV (retrieved date: March 27, 2018).
- [28] Diabetes UK, (2011) Understanding diabetes. Your essential guide.
<https://www.diabetes.org.uk/upload/How%20we%20help/catalogue/Understanding%20diabetes%20LG%20PRINT.pdf> (June 16, 2017).
- [29] National Institute of Adult Continuing Education. (2009) Readability. How to produce clear written materials for a range of readers. <http://shop.niace.org.uk/readability.html> (retrieved date: March 27, 2018).
- [30] Quinn, S., Bond, R., and Nugent, C. (2017) Ontological modelling and rule-based reasoning for the provision of personalized patient education. *Expert Syst.*, 34(2), pp. e12134.
- [31] Tobii Technology AB. (2010) Tobii T/X series Eye Trackers.
<http://www.tobii.com/siteassets/tobii-pro/product-descriptions/tobii-pro-tx-product-description.pdf> (retrieved date: March 27, 2018).
- [32] Olsen, A. (2012) The Tobii I-VT Fixation Filter. Algorithm description.
<https://www.tobii.com/siteassets/tobii-pro/learn-and-support/analyze/how-do-we-classify-eye-movements/tobii-pro-i-vt-fixation-filter.pdf> (retrieved date: March 27, 2018).
- [33] Bull, F. C., Holt, C., Kreuter, M., Clark, E., and Scharff, D. (2001) Understanding the effects of printed health education materials: which features lead to which outcomes? *J. Health Commun.*, 6(3), pp. 265-280.
- [34] Lustria, M. L. A., Cortese, J., Noar, S. M., and Glueckauf, R.L. (2009) Computer-tailored health interventions delivered over the Web: review and analysis of key components. *Patient Educ. Couns.*, 74(2), pp. 156-173.
- [35] Bol, N., van Weert, J. C., Loos, E. F., Romano Bergstrom, J. C., Bolle, S., and Smets, E. M. (2016) How are online health messages processed? Using eye tracking to predict recall of information in younger and older adults. *J. Health Commun.*, 21(4), pp. 387-396.
- [36] Bol, N., van Weert, J. C., de Haes, H. C., Loos, E. F., de Heer, S., Sikkel, D., and Smets, E. M. (2014) Using cognitive and affective illustrations to enhance older adults' website satisfaction and recall of online cancer-related information. *Health Commun.*, 29(7), pp. 678-688.
- [37] Bakhshi, S., Shamma, D. A., and Gilbert, E. (2014) Faces engage us: Photos with faces attract more likes and comments on Instagram. *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, 2014, pp. 965-974.
- [38] Fletcher-Watson, S., Findlay, J. M., Leekam, S.R., and Benson, V. (2008) Rapid detection of person information in a naturalistic scene. *Perception*, 37(4), pp. 571-583.
- [39] Birmingham, E., Bischof, W. F., and Kingstone, A. (2009) Saliency does not account for fixations to eyes within social scenes. *Vision Research*, 49(24), pp. 2992-3000.
- [40] Cerf, M., Harel, J., Einhäuser, W., and Koch, C. (2008) Predicting human gaze using low-level saliency combined with face detection. In Platt, J. C., Koller, D., Singer, Y., and Roweis, S. T. (eds), *Advances in Neural Information Processing Systems 20: 21st Annual Conference on Neural Information Processing Systems 2007*. Curran Associates Inc.
- [41] Tsai, C., and Guan, S. (2015) Identifying regions of interest in reading an image. *Displays*, 39, pp. 33-41.
- [42] Birmingham, E., Bischof, W. F., and Kingstone, A. (2007) Why do we look at people's eyes? *Journal of Eye Movement Research*, 1(1).