Towards a science of user engagement (Position Paper)

Simon Attfield
University of Middlesex
s.attfield@mdx.ac.uk

Gabriella Kazai
Microsoft Research
v-gabkaz@microsoft.com

Mounia Lalmas
Yahoo! Research
mounia@acm.org

Benjamin Piwowarski
University of Glasgow
benjamin@bpiwowar.net

ABSTRACT

User engagement is a key concept in designing user-centred web applications. It refers to the quality of the user experience that emphasises the positive aspects of the interaction, and in particular the phenomena associated with being captivated by technology. This definition is motivated by the observation that successful technologies are not just used, but they are engaged with. Numerous methods have been proposed in the literature to measure engagement, however, little has been done to validate and relate these measures and so provide a firm basis for assessing the quality of the user experience. Engagement is heavily influenced, for example, by the user interface and its associated process flow, the user's context, value system and incentives.

In this paper we propose an approach to relating and developing unified measures of user engagement. Our ultimate aim is to define a framework in which user engagement can be studied, measured, and explained, leading to recommendations and guidelines for user interface and interaction design for front-end web technology. Towards this aim, in this paper, we consider how existing user engagement metrics, web analytics, information retrieval metrics, and measures from immersion in gaming can bring new perspective to defining, measuring and explaining user engagement.

Categories and Subject Descriptors

H.1.1 [**MODELS AND PRINCIPLES**]: User/Machine Systems—*Human factors*

General Terms

Measurement

Keywords

User experience, User engagement, Evaluation

1. INTRODUCTION

Over the last two decades the nature of human computer interaction has been transformed. In the 1980s the dominant model was of a single person sat in front of a workstation, probably in an office. At the end of the day computing was left in the office and interaction came to an end. Since then computers and connectivity have come to pervade all aspects of our lives. We have computers in our homes and in our pockets. The continually evolving web provides unprecedented choice of things we can do; we can educate ourselves or others, buy and sell, watch movies, keep up with friends, write blogs, share content, play games, or simply indulge a passing curiosity. With this the conceptualisation of the user has also changed. Kuutti [23] describes this evolution as being from a cog in an organisational machine and a "source of error", to a partner in social interaction and ultimately consumer, and more recently, with Web 2.0, to content creator.

In line with this transformation, researchers and practitioners in user-centred design have increasingly seen the need to extend traditional performance-based notions of usability, such as effectiveness, efficiency and satisfaction (ISO 9241-11), to encompass non-utilitarian aspects of interaction (e.g., [12, 24]). Under the broad rubric of user experience, there is emphasis on understanding and designing for the subjective aspects of technology encounters (e.g., [24, 47]). In a world full of choices where the fleeting attention of the user becomes a prime resource, it is essential that technology providers do not just design systems but that they design engaging experiences [31]. Indeed, in an environment of choice, failing to engage the user "equates with no sale on an electronic commerce site and no transmission of information from a website; people go elsewhere" [28]. The question then becomes: how do we do this, and closely related to that, how do we *assess* the experience as being the kind we would like to design?

The answer lies in the study of user engagement, a quality of user experience that emphasises the positive aspects of interaction, and in particular the phenomena associated with being captivated by technology (and so being motivated to use it) [28]. Successful technologies are not just used, they are engaged with; users invest time, attention, and emotion into the equation.

To promote user engagement we need to be able to design for it. To know when we have done this we need some way of assessing and measuring it. Only by knowing when and why engagement occurs can we understand what is effective. As discussed in this paper, many works on user engagement have been written, but they are spread over a range of different areas of research, thus hindering the possibility of building upon previous research. We thus need a framework in which user engagement can be studied, measured and explained, leading to, for example, recommendations and

guidelines for user interface and interaction design, that is, we need a science of user engagement.

Towards this goal, this paper contains our proposal to establish the "science of user engagement". It has two main contributions. First, we provide a review and classification of prior research on user engagement and associated measures (Section 2), with a special focus on the work carried out outside large-scale web log analysis. Based on this review, and our insights, we then present our proposed research agenda providing a detailed roadmap on how to define and set up a framework in which user engagement can be studied, measured, and explained (Section 3).

2. USER ENGAGEMENT IN THE LITERATURE

In this section, we provide a review of prior research on user engagement. We start with our own definition based on our review of that literature.

> User engagement is the emotional, cognitive and behavioural connection that exists, at any point in time and possibly over time, between a user and a resource.

This definition is intentionally broad. By identifying emotional, cognitive, and behavioural factors, it emphasises the holistic character of user engagement and is also suggestive of aspects that are open to measurement. It also refers equally to user engagement in terms of a single session or a more long-term relationship across multiple session. User engagement with a technological resource is not just about how a single interaction unfolds, but about how and why people develop a relationship with technology and integrate it into their lives.

2.1 Characteristics

In the following we discuss some characteristics associated with user engagement, either as presented by previous studies or as suggested by us. These elaborate the notion of engagement over the three broad dimensions: emotional, cognitive and behavioural. While some of the characteristics have stronger ties with one of the dimensions, most are a combination of the three. They also provide some orientation for thinking about the designable causes and observable consequences of user engagement.

Focused attention.
Being engaged in an experience involves focusing attention to the exclusion of other things, including other people (unless social interaction enhances the engagement) [28]. This phenomenon relates to distortions in the subjective perception of time during interaction [29], which has been shown to be an effective indicator of cognitive involvement [2] and has been used as a component in measures of games immersion [18]. The more engaged someone is, the more likely they are to underestimate the passage of time. Together with concentration, absorption and loss of self-consciousness, distortions in the subjective perception of time have led to parallels being drawn between engagement and the idea of flow as an optimal experience [9], where flow refers to a mental state in which a person is fully immersed in what they are doing.

Positive Affect.
"Engaged users are affectively involved" [28]. Affect relates to the emotions experienced during interaction. For example, O'Brian & Toms [28] found that a lack of fun can act as a barrier to shopping online and that fun during a webcast can draw the user in, although a lack of awareness of other viewers can mitigate against the quality of the fun experience. Jennings [19] argues that affective experiences are intrinsically motivating and that, in relation to engagement on the web, an initial affective hook can induce a desire for exploration or active discovery. This, they argue then encourages greater emotional involvement and contributes to customer loyalty (see endurability below). They regard affect as "an emotional investment that helps create a personal link to an experience or activity" [19].

Aesthetics.
Aesthetics concerns the sensory, visual appeal of an interface and is seen as an important factor for engagement [29]. In the context of online shopping, web searching, educational webcasting and video games, O'Brian & Toms [28] relate aesthetics to factors such as screen layout, graphics and the use of design principles such as symmetry and balance. In the context of multimedia design, Jennings [19] relates aesthetics (in this case media quality) to positive affect and suggests that aesthetic experiences promote focused attention and stimulate curiosity. The significance of aesthetics was well demonstrated by Tractinsky & Ikar [45] who, in an experiment in which users interacted with a computerised, surrogate Automated Teller Machine, found a positive correlation between perceived usability and aesthetic appeal, even though no such correlation existed between perceived usability and actual usability.

Endurability.
People remember enjoyable, useful, engaging experiences and want to repeat them. This aspect of engagement refers to the likelihood of remembering an experience and the willingness to repeat it [35]. It also relates to a users' willingness to recommend an experience to others, and to their perceptions of whether an experience met their expectations of being "successful", "rewarding", or "worthwhile" [29]. In the context of the web, endurability can be related to the notion of 'sticky' content and the goal of holding the user's attention and encouraging them to return. O'Brian & Toms [28] found that having fun, being rewarded with convenience and incentives, and discovering something new promoted later re-engagement.

Novelty.
Interactive experiences can be engaging because they present users with novel, surprising, unfamiliar or unexpected experiences. Novelty appeals to our sense of curiosity, encourages inquisitive behaviour and promotes repeated engagement [29]. It can arise through freshness of content or innovation in information technology [13]. In e-commerce applications, shoppers can enjoy the experience of becoming sidetracked, browsing just to see what is there [28]. It has been shown that learners experience higher levels of engagement during multimedia presentations that exhibit higher levels of variety [48]. However, there is also suggestion of the importance of a subtle balance between novelty on the one hand and familiarity on the other. For example, in gam-

Characteristic	Definition	Measures	Ref.
Focused Attention	Focusing attention to the exclusion of other things	Distorted perception of time, follow-on task performance, eye tracking	[28, 29, 2, 18, 9, 15]
Positive Affect	Emotions experienced during interaction	Physiological sensors (e.g. face detection)	[28, 19, 15]
Aesthetics	Sensory and visual appeal of an interface	Online activity (curiosity-driven behaviour), Physiological sensors (e.g., eye tracking), perceived utility	[29, 28, 19, 45, 15]
Endurability	Likelihood of remembering an experience and the willingness to repeat or recommend it	Online activity (e.g. bookmarking, sending emails)	[35, 29, 28, 32, 34, 49]
Novelty	Novel, surprising, unfamiliar or unexpected experiences	Physiological sensors (e.g., blood pressure)	[29, 13, 28, 48, 39, 43]
Richness and Control	Levels of richness and control	Online activity (e.g., interaction with the site, time spent), Physiological sensors (e.g. mouse pressure)	[37, 32, 46]
Reputation, trust and expectation	Global trust users have on a given entity	Online activity (returning user, recommendation)	[22, 27, 44, 10, 20, 41]
User Context	User's motivation, incentives, and benefits	Online activity (location, time, past history)	[26, 25, 30, 11]

Table 1: In this table, we summarise the identified characteristics of user engagement discussed in Section 2, give the possible ways to objectively measure them (beyond questionnaires, if any), and give the references where such characteristics are discussed

ing the level of novelty can determine whether engagement is sustained [28], although some familiarity with a game environment can lead to faster engagement [39] and reduced disorientation [48].

Richness and control.
The "Richness, Control and Engagement" (RC & E) framework [37] explains levels of engagement in terms of the levels of richness and control that are shaped by the features of a product and the user's expertise. Richness captures the growth potential of an activity by assessing the variety and complexity of thoughts, actions and perceptions as evoked during the activity (e.g., variety, possibilities, enjoyment, excitement, challenge). Control captures the extent to which a person is able to achieve this growth potential by assessing the effort in the selection and attainment of goals (clarity, ease, self confidence, freedom). Through experiments on a digital voicemail system, where they varied the number of features of the user interface, the amount of voicemail content and the type of task, Rozendaal et al. [37] showed that level of engagement could be predicted according to the level of richness and control experienced.

Reputation, trust and expectation.
Trust is a necessary condition of user engagement. Reputation can be seen as the trust users invest globally in a given resource or provider. Trust is not only a matter of the insurances offered by technology (e.g. encrypted communications), but also depends on implicit contracts between people, computers, and organisations [22]. Organisations in "real life" do rely on consumers' faith in the information or service provided. For example, telephone companies are required to correctly connect calls and protect the integrity of phone numbers. Similar requirements can be issued when transposing trust to the web. Where users interact with each other, the reputation of the web site has a strong influence

on the extent to which users trust each other in their transactions, e.g., on eBay [27]. Similarly, the perceived fairness of the scoring system in Q&A sites influences users' trust in the service and their willingness to engage [44]. In social networks, trust naturally has an additional dimension related to the perception of trustworthiness of user-generated content [10]. Trust can also be related to engagement with web search services, where authoritativeness and popularity features are used to rank search results [20]. Finally, a consequence of trust or reputation is expectation, which harnesses engagement before a user has even reached a web site [41].

User context.
User's motivation, incentives, and benefits affect the experience more than in traditional usability [26]. User experience is also very context dependent [25], so the experience with the same design in different circumstances is often very different. This means that user engagement evaluation cannot be conducted just by observing user's task completion in a laboratory test [30]. Similarly, the range of available choices, as well as the accepted social norms, values and trends impact on how users engage [11]. The user's personal preferences and priorities over aspects that influence engagement, such as trendiness, coolness, or fun are likely to change in different usage scenarios and domains. In addition, different forms of engagement are likely to suit different types of personality (e.g., couch potatoes, critics, or creators).

These characteristics go some way to elaborate the emotional, cognitive and behavioural components of user engagement. They also say something about how engagement might be recognised, and even promoted. However, research is still very patchy. We know little about how different characteristics might be more or less significant for different scenarios of web interaction and different user demographics.

There may be more significant characteristics yet to be identified. Researchers are only beginning to explore this space. We return to this in Section 3. For now these characteristics provide some initial leverage on approaches to measuring user engagement. In the next section we explore some of these opportunities.

2.2 Measurement

Having defined user engagement and elaborated some of its key characteristics we now look into potential approaches to its assessment. Since user engagement is multi-faceted (as witnessed by the list above) there are (and should be) many approaches to its measurement.

User experience evaluation metrics can be divided into two broad types: subjective and objective. Subjective measures record a user's perception, generally self reported, of the media at hand. User's subjective experiences are central to user engagement and we consider methods for assessing these. Subjective experiences, however, can have objectively observable consequences, and so we also consider objective measurements that may be indicative of user engagement. These include independent measures such as the passage of time or number of mouse clicks to complete a task.

2.2.1 Subjective measures

Where the subjective characteristics of an interactive experience have been defined, it is possible to construct a post-experience questionnaire to measure them in relation to a given interactive experience. Such an instrument was developed by O'Brian, Toms and colleagues. They first [28] report a literature review and exploratory interview study with users in four domains (online shopping, web searching, educational webcasting, and video games). Through their analysis they identify a set of attributes of user engagement (some of which are listed above). They then used this initial conceptualisation to develop a questionnaire to assess user-engagement among online shoppers [29]. They administered their survey to 440 online shoppers and used factor analysis to reduce the characteristics to a set of key constructs. Their survey provides a standardised instrument for eliciting users engagement assessments.

O'Brian et al's aim was to produce a general purpose user engagement questionnaire. However, user engagement almost certainly has different characteristics in different application domains and for different demographic groups. The web offers a diversity of experiences relevant to a diversity of users. The properties of engagement during instant messaging, for example, may differ from the properties of engagement with a news portal. Aesthetic appeal may be an important factor for engagement with a films website whilst trust has been found to be a key factor in engagement with health websites [50]. Related to this, different user populations may have different priorities. Whereas fun may be an important characteristic for engaging children [35], ease of navigation may be a higher priority for adults. Some characteristics of engagement may generalise well, others may not.

This raises the question of generating new user engagement instruments relevant to different application domains and user groups. As O'Brian and Toms demonstrate, an approach is to begin with exploratory, qualitative studies designed to uncover engagement characteristics for specific kinds of interaction and user. Analysis can then draw out key variables indicative of engagement in this context, followed by the creation and validation of questionnaires that probe on these variables.

One issue that arises with post-experience questionnaires and tests, however, is that they are not sensitive to ways in which an interaction changes over time. Interactions are dynamic and engagement fluctuates [28]. In the interests of differentiating aspects of an interactive experience it may be valuable to measure such temporal changes. In this case post-experience questionnaires may not be the best tool, and objective measures seem more well suited.

2.2.2 Objective measures

There are a number of drawbacks to questionnaires and other subjective measures. These include their reliance on the user's subjectivity [42], post-hoc interpretation and their susceptibility to the halo effect[1] [38]. A common strategy to overcome these is to develop objective measures that can reliably indicate subjective states [16]. This is done, for example, in studies of presence [40] and games immersion [18].

This raises the question of what objective phenomena are indicative of engagement. To date little has been reported on this, but since engagement is multifaceted and complex, a number of different variables may be useful to measure. In the following we discuss some measures used in previous research. What becomes evident is that each objective measure tend to target a very specific aspect of engagement, unlike questionnaires, which can address a range of variables. In light of this and our proposal that different engagement characteristics are relevant for different kinds of situation, we anticipate that it will be both useful and necessary to establish a range of objective measures and use this range to predict engagement.

The measures we consider here are: the subjective perception of time, follow-on task performance, physiological sensors, online behaviour, and information retrieval metrics. In Table 1, we suggest how these measures could be related to different aspect of user engagement.

The subjective perception of time (SPT).
Assessing the subjective perception of time involves asking a user to make some estimation of the passage of time during an activity. This can involve retrospectively probing for how long an activity lasted or asking a user to indicate the duration of fixed time-intervals during the activity [2]. Despite its name, we regard SPT as an objective measure given that it is assessed against actual time. To ask somebody to report how long an activity has lasted or estimate an interval is to ask them a question they can get wrong (by different degrees). In a sense it is a measure of performance. As discussed above, SPT has been used in attention research and is indicative of cognitive aspects of engagement.

Follow-on task performance.
Another potential measure of cognitive engagement is how well somebody performs on a different task immediately following a period of engaged interaction. Games researchers have found that the more immersed a person is when playing

---

[1]This a cognitive bias whereby the perception of one trait (i.e. a characteristic of a person or object) is influenced by the positive evaluation of another trait (or several traits) of that person or object.

a game, the longer it takes them to complete an unrelated puzzle task immediately afterwards [18]. A related measure is to assess how well somebody performs on a secondary task that periodically interrupts an engaged interaction.

Physiological measures.
Physiological data can be captured by a broad range of sensors [15] related to different cognitive states. Examples of sensors are eye trackers (difficulty, attention, fatigue, mental activity, strong emotion), mouse pressure (stress, certainty of response), biosensors (e.g. temperature for negative affect and relaxation, electrodermal for arousal, blood flow for stress and emotion intensity), oximiters (e.g., pulse), camera (e.g., face tracking for general emotion detection). Such sensors have several advantages over questionnaires or online behaviour, since they are more directly connected to the emotional state of the user, are more objective (involuntary body responses) and they are continuously measured. They are however more invasive and cannot be used in large scale studies.

Ikehara and Crosby [15] have used such sensors to assess the cognitive load of a "game" (tracking fractions). Jennett et al. [18] also report the use eye-tracking data to assess immersion. They found that eye-movement increases over time during a non-immersed experience and reduces over time during an immersed experience. Users seem to focus on fewer targets during engaged attention.

In general, such measures could be highly indicative of immersive states through their links with attention, affect, the perception of aesthetics and novelty. Research is required to refine our understanding of precisely what physiological states are indicative of engagement and how to differentiate against negative states of high arousal such as stress.

Online behaviour.
The subjective perception of time, follow-on task performance and physiological indicators are objective measures that are potentially suitable for measuring a small number of interaction episodes at close quarters. In contrast, the web-analytics community has shown some interest in measuring user engagement through various approaches that assess users' depth of engagement with a site. In this context Peterson [32] defines engagement as, "An estimate of the degree and depth of visitor interaction against a clearly defined set of goals." For example, Peterson [32] considers measures that indicate visitors who consume content slowly and methodically, return directly to a site, and whether they subscribe to feeds, and BusinessWeek.com have defined a user-engagement metric calculated on the basis of the number of comments per posting on their site [46].

So-called interaction patterns, like for example web search logs, can be instrumental in studying user engagement. A step in this direction can be found in [34] who proposed to summarise a series of actions performed by the user by a single value. A related approach can be found in the patent Methods and systems for detecting user satisfaction (US Patent 7587324) where the inventors propose to measure user satisfaction by assigning a utility value to each sequence of actions (between a user and a computer interface) and comparing this value to a baseline. User interaction patterns have also be used to predict search engine switching behaviour [49], which is a facet of user engagement.

From a commercial perspective, assessments of online be-

haviour fall within the scope of customer engagement (CE). CE is an instrument of marketing referring to the engagement of customers with one another, with a company or a brand. Its more recent form considers online CE. Online CE enables organisations to respond to the changes in customer behaviour that the internet has brought about, as well as the increasing ineffectiveness of the traditional "interrupt and repeat" broadcast model of advertising. Supporting "deep" engagement is seen as an important source of competitive advantage, whether through advertising, user generated product reviews, FAQs, forums where consumers can socialise or contribute to product development. In this space, a number of initiatives have emerged to address the question of CE metrics from organisations such as the World Federation of Advertisers (WFA), the Association of National Advertisers (ANA) and Nielsen Media Research. As in the academic literature, engagement is seen as multifaceted and open to a range of potential metrics [43].

Information retrieval (IR) metrics.
Among the different lines of research in IR metrics, three are directly related to measuring user engagement. The first is to develop metrics for interactive IR, e.g., [5, 17] where users and their contexts are taken into account. This line has brought up the idea of simulated search scenarios, where a subject is asked to follow a search scenario that specifies what, why, and in which context the user is searching.

The second line of research is to develop metrics that incorporate enriched user interaction models, e.g., [33, 21, 51]. This is needed when old assumptions typical of traditional IR (linear browsing, binary relevance) do not hold. The work of e.g. Piwowarski & Dupret [33] and Kazai et al. [21] provides formal user models that could be adapted to defining the behaviour of a user when interacting with a web service, and hence a way to provide measures correlated with user engagement.

Finally, the third line of research relates standard IR effectiveness metrics to user satisfaction, e.g., [1, 14]. The research of Al-Maskari et al. [1] shows that we can expect some correlation between carefully chosen metrics and user-oriented measures, even with metrics that were not specifically designed to capture such aspects of the retrieval process. However, more research is needed in order to validate and extend these results.

2.3 Building on Foundations

Interest is growing in user engagement and its measurement, as well as that of related concepts. However work is somewhat fragmented, even if more holistic frameworks begin to appear [36]. This framework has been developed for defining large-scale user-centred metrics, both attitudinal and behavioural, coined HEART (Happiness, Engagement, Adoption, Retention, and Task success), and relating it to product goals. Our approach is broader and more systematic, in the sense that we want to capture and model users, and not only design user metrics, and relate these models and measures to design principles for web applications.

It is unclear exactly what to measure and how to measure it and what the important measures are for specific scenarios. It is also unclear how different aspects of user engagement relate to each other in these scenarios. A main observation is that there seems to be a lack of understanding of how to integrate subjective and objective measures.

Note, for the former, experiments were specifically designed to study user engagement or something related to it. And whilst the latter hold the promise of being able to assess dynamic fluctuations in engagement, precise correlations with engagement are yet to be established.

An important next step is to look carefully at the work described above, and explore the possible mappings between subjective and objective variables in specific contexts. Taking into account a range of interaction contexts adds to the complexity of this task, but is imperative for enabling progress. Experiments will have to be carefully designed, and validation will need to be large-scale. Here we have addressed a number of approaches with potential for making progress in this area. Whilst new ground needs to be made, however, our review outlines a solid foundation of existing work in HCI, IR and web evaluation which can provide the basis for this work. Importantly, we aim to integrate mappings as empirically grounded models. In this way we want to move towards a "science" of user engagement that can have fundamental impact in how to design the user experience.

3. APPROACH

Our goal is to define a framework in which user engagement can be studied, measured, and explained, and, as an ultimate aim, lead to recommendations and guidelines for user interface and interaction design for front-end web technology.

Our approach is holistic but convergent, following a general progression from: (i) Empirical and analytic exploration of the characteristics of user engagement (ii) Correlating and resolving these to integrated models and reliable measures both objective and subjective (iii) Design and validation of prototypes

Within the approach, we propose the following three main directions of work: (1) identification of interaction patterns and development of engagement measures for a range of contexts and services; (2) adaptation of immersion concepts from gaming; and (3) designing for user engagement. The first two lines of research try to define user engagement in various contexts, along with metrics and methodologies for measuring it. The third gives direct and practical means to deploying definitions, metrics and methodologies for designing engaging technologies.

These lines of research are not independent from each other, but represent related aspects that will undoubtedly influence and inform the overall approach for defining and measuring user engagement. We elaborate on each in the next sections.

3.1 Developing measures and models of user engagement

The goal of the first line of investigation is to obtain an understanding of how a range of user engagement characteristics relate to specific user contexts and scenarios and how they can be calculated using state-of-the-art web analytic metrics, IR metrics, and existing or novel user engagement measures.

This would expand on the work of O'Brian and colleagues at the University of Dalhousie [28, 29]. In our opinion, their work presents the most significant progress towards understanding and assessing user engagement to date. However, whereas the Dalhousie group assumes generality across ac-

tivity domains and user groups, an important goal will be to derive reliable and reusable metrics for user engagement that are specific to particular contexts and online services.

In order to do so, it is necessary to collect and catalogue existing and new engagement characteristics and indicators. Such characteristics can be tested for correlation with a range of existing measures, such as standard IR metrics for assessing web search engines and web analytics measures. This would involve setting up a series of "lab-based" user experiments (combinations of virtual, online lab experiments, actual lab environments, and in-situ studies) around specific user tasks and usage scenarios in which users' perceived values are studied for various characteristics through questionnaires and other subjective measures. Data would also be collected through observation and logging.

It is important that experiments are carried out on a sufficiently large and diverse sample of the user population and using a (feasibly) wide range of design variants, user tasks and motivations to ensure generalisability. To reduce bias, appropriate sampling methods with respect to the user population have to be used. Psychological techniques such as experience sampling[2] (e.g., [8]) can be used to gain insights on user interactions.

From these experiments, dependency relationships can be established among the different engagement characteristics, leading to classes of user models for the different contexts and task types. A specific point of interest is the relationship between incentive structures, rewards, benefits and engagement. The derived user engagement models can then be validated through further experiments.

In order to assess the full range of variables of engagement, it is necessary to test a multitude of scenarios, interaction and graphics designs, for example, different page layouts, result presentations, modified user interaction models (such as result peek, scratch pad), etc. We anticipate the need for continuous refinement to this methodology, allowing fluid adaptation to the latest findings and the research questions that such results raise.

Summarising, this line of research will deal with: (1) Classifying characteristics of user engagement as relevant to particular contexts, user tasks, and incentives, and in particular, validating characteristics relevant to web interactions (extending our current work); (2) Identifying whether characteristics can be automatically derived using standard or more advanced web analytic measures or IR metrics and developing prediction models when applicable; (3) Identifying patterns of user engagement (leading to the development of models of user engagement), both from qualitative (surveys, questionnaires, focus groups, participatory design, etc.) and quantitative points of view (through log analysis); (4) Developing from this a methodology for measuring and understanding user engagement, including methods for comparative measures across brands and methods for failure analysis.

3.2 New perspectives from immersion in gaming and related areas

The goal of the second line of research is to develop novel concepts of user engagement through an examination of the notion of immersion as it is used in gaming. Being immersed

---

[2]Essentially, participants fill out several brief questionnaires every day by responding to alerts, where the questionnaires ask about their current activities and feelings.

in a computer game involves being "drawn in", with attention focussed entirely on the game [18]. As such, this phenomenon has much in common with engagement and many ideas emerging from this and related research may be applicable to more general notions of user engagement and notions of user engagement on the web in particular. Some early work on the application of the concept of immersion is Bhatt's [3] preliminary study of interaction with a range of e-commerce web sites. Bhatt finds the relative importance of immersion (compared to interactivity and connectivity) to be genre-dependent; for example, it is far more significant to the fashion industry than it is to the financial industry. This finding concurs with our initial view of the applicability of immersion to user engagement and also the context dependant nature of user engagement.

In relation to immersion, of particular interest is the work of Cairns and colleagues. Their early work [6] reported a user study of immersion in gaming. They observed three progressively deeper levels of immersion, which they referred to as engagement, engrossment, and total immersion. In a more recent paper [7] they review the general notion of user experience in the light of their Core Elements of the Gaming Experience (CEGE) model. CEGE is a hierarchical model that elaborates a set of necessary conditions of an enjoyable gaming experience. For example, a positive experience depends on the user taking "ownership" of a game, where ownership involves developing high-level strategies and goals and receiving rewards. But this itself is built upon lower-level factors such as "control" elements (e.g., interface factors, small actions, point-of-view) and "facilitators" (such as time to play and the user's aesthetic values). Calvillo-Gamez et al. [7] argue that many of these factors can be meaningfully extrapolated to other kinds of interaction, such as the use of productivity applications.

We propose further exploration of the concept of immersion and concepts from related domains (such as flow, and presence in Virtual Reality experiences) in order to expand and enrich existing notions of engagement. Underpinning this work will be the aim of expanding our understanding of the necessary constituents of an engaging experience. We believe that this can bring valuable new perspectives to the measurement of user engagement on the web.

The expected outcomes are: (1) Understanding the links (and contrasts) between user engagement and immersion (and to some extent flow and presence); (2) Leveraging this understanding to enrich our model of user engagement; (3) Developing additional metrics relevant to search and different genres of web interaction; (4) Conceptual input into larger-scale experimentation in tandem with the first line of work (Section 3.1).

3.3 Designing for user engagement

This last line of research explicitly focuses on designing for user engagement, by providing a useful source of validated design ideas. This is based on insights from the other two lines of research to inform the development, validation and refinement of design concepts (initiated in the first line of research – Section 3.1) for services and tools aimed at promoting user engagement.

To conduct such a research, an iterative, user-centred design process incorporating evaluation throughout, should be adopted. Initial qualitative field studies can be used to identify and characterise particular kinds of interaction scenario and their requirements. Scenarios should include individual interaction as well as collaboration, sharing and social interactions. Each scenario would provide a basis for developing conceptual designs, which implement generalisable strategies for promoting engagement appropriate to the context.

Prototypes ranging from low to high-fidelity can be developed and evaluated iteratively, with early formative evaluations leading to the development of design improvements. Summative evaluations can then be used to assess design alternatives in terms of user engagement metrics developed in other areas of the work. Overall the process will be aimed at generating and testing generalisable design strategies for promoting engagement in different contexts, ultimately resulting in reusable recommendations, guidelines and exemplars which can be employed by other designers.

An approach to the cataloguing and communication of design strategies, which we favour, is to record these in the form of interaction design patterns [4]. Patterns are an approach to capturing reusable solutions to recurring design problems within semi-formal representations. As such they offer a convenient basis for generating a 'lingua franca' of design solutions that encapsulate a particular set of design values (in this case the promotion of engagement).

The expected outcomes of this line of research are to: (1) Evaluate prototype solutions for promoting user engagement in a range of contexts; (2) Produce generalisable design recommendations in the form of reusable interaction design patterns.

4. CONCLUSIONS

In this paper, we advocate the development of an approach for studying, measuring, and explaining user engagement. We first define user engagement, and then report on and classify various characteristics and potential methods for measurement. User engagement is a multifaceted, and complex phenomenon; this applies to both its definition and empirically grounded "signature" characteristics. Both of these give rise to a number of potential approaches for measurement whether objectively or subjectively.

We proposed three lines of research that have the potential to shape what we like to refer as the science of user engagement. The first two define methodologies for measurement, whilst the third outlines how these methodologies can lead to user engagement aware front-end technology design. Our proposed approach is motivated by the aim of achieving an innovative yet principled approach to the problems associated with measuring user engagement, building on existing solid and fundamental work and methodology, and taking inspiration from other related areas.

5. REFERENCES

[1] A. Al-Maskari, M. Sanderson & P. Clough. The relationship between IR effectiveness measures and user satisfaction. ACM SIGIR, 2007.

[2] D. Baldauf, E. Burgarda & M. Wittmann. Time perception as a workload measure in simulated car driving. Applied Ergonomics, 40(5):929-935, 2009.

[3] G. Bhatt. Bringing virtual reality for commercial web sites. Human-Computer Studies, 60(1):1-15, 2004

[4] J. O. Borchers. A pattern approach to interaction design. AI & Society, 15(4): 359-376 (2001)

[5] P. Borlund & I. Ruthven. Introduction to the special issue on evaluating interactive information retrieval systems. IP&M, 44(1): 1-3 (2008)

[6] E. Brown & P. Cairns. A grounded investigation of immersion in games. ACM CHI, pp 1297-1300, 2004.

[7] E. Calvillo-Gamez, P. Cairns & A. Cox. From the Gaming Experience to the Wider User Experience. BCS HCI 2009, pp 520-523, 2009.

[8] S. Consolvo & M. Walker. Using the experience sampling method to evaluate ubicomp applications.

[9] M. Csikszentmihalyi. Flow: The psychology of optimal experience. Harper and Row, 1990.

[10] J. Golbeck. Computer Science: Weaving a Web of Trust. Science, 321(5896):1640-1641, 2008.

[11] R. Harper, T. Rodden, Y. Rogers & A. Sellen. Being Human: Human-Computer Interaction in the Year 2020. Microsoft, Cambridge, UK. 2008.

[12] M. Hassenzahl & N. Tractinsky. User experience - a research agenda. Behaviour and Information Technology, 25(2):91-97, 2006.

[13] M. Huang. Designing website attributes to induce experiential encounters. Computers in Human Behavior, 19(4):425-442, 2003.

[14] S. B. Huffman & M. Hochster. How well does result relevance predict session satisfaction?. ACM SIGIR, 2007.

[15] C. S. Ikehara and M. E. Crosby. Assessing cognitive load with physiological sensors. In Proceedings of the 38th Annual Hawaii International Conference on System Sciences, Washington, DC, USA, 2005. IEEE Computer Society.

[16] W.A. Ijsselstein, H. de Ridder, J. Freeman & S.E. Avons. Presence: Concept, determinants and measurement. SPIE, 2000.

[17] P. Ingwersen & K. Jarvelin. The Turn: Integration of Information Seeking and Retrieval in Context. Springer-Verlag, 2005.

[18] C. Jennett, A. Cox, P. Cairns, A. Dhoparee, A. Epps, T. Tijs & A. Walton. Measuring and Defining the Experience of Immersion in Games. International Journal of Human-Computer Studies 66(9):641-661, 2008.

[19] M. Jennings. Theory and models for creating engaging and immersive e-commerce websites. ACM SIGCPR, 2000.

[20] G. Kazai & N. Milic-Frayling: Effects of Social Approval Votes on Search Performance. ITNG, 2009.

[21] G. Kazai, N. Milic-Frayling & J. Costello. Towards methods for the collective gathering and quality control of relevance assessments. ACM SIGIR, 2009.

[22] R. Khare & A. Rifkin. Weaving a Web of trust. World Wide Web J., 2(3):77-112, 1997.

[23] K. Kuuttii. Hunting for the lost user: From sources of errors to active actors and beyond. Cultural usability seminar, University of Art and Design Helsinki, 2001.

[24] E. L-C. Law, N. Bevan, G., Christos, M. Springett & M. Larusdottir, (Eds.). Proceedings of the COST294-MAUSE Workshop on Meaningful Measures: Valid Useful User Experience Measurements (VUUM), 2008.

[25] E. Law, V. Roto, M. Hassenzhal, A Vermeeren & A. Kort. Understanding, scoping and defining user eXperience: A survey approach. ACM CHI, 2009.

[26] A. Mäkelä & J. Fulton Suri. Supporting Users' Creativity: Design to Induce Pleasurable Experiences. International Conference on Affective Human Factors Design, 2001.

[27] M. M. L. Mui. A Computational Model of Trust and Reputation for E-businesses. HICSS - Volume 7, 2002.

[28] H. O'Brian & E. Toms. What is user engagement? A conceptual framework for defining user engagement with technology. JASIST, 59(6):938-955, 2008.

[29] H. O'Brian, E. Toms, K. Kelloway & E. Kelley. The development and evaluation of a survey to measure user engagement. JASIST 61(1):50-69 (2010)

[30] M. Obrist, V. Roto, & K. Väänänen-Vainio-Mattila. User experience evaluation: do you know which method to use? CHI Extended Abstracts, 2009.

[31] K. Overbeeke, T. Djajadiningrat, C. Hummels, S. Wensveen & J. Frens. Lets make things engaging. Funology. Kluwer, 2003.

[32] E.T. Peterson. How do you calculate engagement? Part I. Web Analytics Demistified (blog) http://blog.webanalyticsdemystified.com/weblog/2006/12/how-do-you-calculate-engagement-part-i.html (accessed Oct 2009), 2006.

[33] B. Piwowarski & G. Dupret. Evaluation in (XML) information retrieval: expected precision-recall with user modelling (EPRUM). ACM SIGIR, 260-267, 2006.

[34] B. Piwowarski, G. Dupret & R. Jones. Mining user web search activity with layered Bayesian networks or how to capture a click in its context. ACM WSDM, 2009.

[35] J. Read, S. MacFarlane & C. Casey. Endurability, engagement and expectations: Measuring children's fun. Interaction Design and Children (2):1-23, 2002.

[36] K. Rodden, H. Hutchinson, and X. Fu & Measuring the User Experience on a Large Scale: User-Centered Metrics for Web Applications. Proceedings of CHI, 2010.

[37] M. Rozendaal, D.V. Keyson & H. Ridder. Product features and task effects on experienced richness, control and engagement in voicemail browsing. Personal Ubiquitous Computing, 13(5):343-354, 2009.

[38] F. Saal, R. Downey & M. Lahey. Rating the ratings: assessing the psychometric quality of rating data. Psychological Bulletin, 88 (2):413-428, 1980.

[39] N. Said. An engaging multimedia design model, Interaction Design and Children. Interaction design and children: building a community, pp 169-172, 2004.

[40] M.V. Sanchez-Vives & M. Slater. From presence to consciousness through virtual reality. Nature Reviews Neuroscience 6(4):332-339, 2005.

[41] J. Saipe. User Experience (UX) and Improving Engagement. E-insight Blog: Usability & UX. http://www.emarketeers.com/e-insight/usability-ux/user-experience-ux-and-improving-engagement.php. 2009.

[42] M. Slater. Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. Presence 8(5). 1999.

[43] P. Summerfield. Getting to engagement: Media agencies start to realize it's part art, part science. Strategy Magazine, June, 2006.

[44] C. Tong Lee, E. Mendes Rodrigues, G. Kazai, N. Milic-Frayling & A. Ignjatovic. Model for Voter Scoring and Best Answer Selection in Community Q&A Services. WI, pp. 116-123. 2009.

[45] N. Tractinsky, A.S. Katz & D. Ikar. What is beautiful is usable. Interacting with Computers (13):127-145, 2000.

[46] E. Ulken. Measuring user engagement: Lessons from BusinessWeek. Online Journalism Review http://www.ojr.org/ojr/people/eulken/200904/1696/ (accessed Oct 2009), 2009.

[47] K. Väänänen-Vainio-Mattila, V. Roto & M. Hassenzahl. Now Let's Do It in Practice: User Experience Evaluation Methods in Product Development. CHI, 2008.

[48] J. Webster & H. Ho. Audience engagement in multimedia presentations ACM SIGMIS, 28(2):63-77, 1997.

[49] R.W. White & S. T. Dumais. Characterizing and predicting search engine switching behavior. ACM CIKM, 2009.

[50] J. Wortham. The Blog That Ignited a Privacy Debate on Facebook, New York Times. February 18, 2009.

[51] E. Yilmaz, M. Shokouhi, N. Craswell & S. Robertson, Incorporating user behavior information in IR evaluation. SIGIR UUIR workshop, 2009.