Rethinking Engagement with Online News through Social and Visual Co-Annotation

Gavin Wood 1, Kiel Long 1, Tom Feltwell 1, Scarlett Rowland 1, Phillip Brooker 3, Jamie Mahoney 1, John Vines 1, Julie Barnett 2, Shaun Lawson 1

1 Northumbria University, Newcastle upon Tyne, UK, {gavin.wood, tom.feltwell, scarlett.rowland, jamie.mahoney, john.vines, shaun.lawson}@northumbria.ac.uk
2 Department of Psychology, University of Bath, Bath, UK, j.e.barnett@bath.ac.uk
3 University of Liverpool, Liverpool, UK, p.d.brooker@liverpool.ac.uk

ABSTRACT
The emergence of fake news, as well as filter bubbles and echo chambers, has precipitated renewed attention upon the ways in which news is consumed, shared and reflected and commented upon. While online news comments sections offer space for pluralist and critical discussion, studies suggest that this rarely occurs. Motivated by common practices of annotating, defacing and scribbling on physical newspapers, we built a mobile app – Newsr – that supports co-annotation, in the form of graffiti, on online news articles, which we evaluated in-the-wild for one month. We report on how the app encouraged participants to reflect upon the act of choosing news stories, whilst promoting exploration, the critique of content, and the exposure of bias within the writing. Our findings highlight how the re-design of interactive online news experiences can facilitate more directed, “in-the-moment” critique of online news stories as well as encourage readers to expand the range of news content they read.

Author Keywords
Online news; social annotation; graffiti.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Engaging with the news is increasingly a social experience that is often entwined with social media practices [22]. Users of almost every social media platform, for instance, routinely share links to news articles and, reciprocally, seek further articles that have been recommended, endorsed or shared by friends; indeed, news recommendation systems themselves utilise our friends’ choices [27]. News websites also typically provide “below the line” (BTL) spaces for readers to comment on professionally-authored articles, whilst many media outlets again use social media platforms to repost links to articles and promote sharing and commentary. At the same time, it has been noted that online sharing and consumption of news in this way can create filter bubbles and echo chambers [11] where users are, for the most part, provided with news and opinions that resonate with their own. It has been noted that BTL sections, in particular, can reinforce discussion of negative portrayals of the people and events that are the subject of the news articles [16]. This is despite research that has also highlighted the potential of public commentary features to provide an online space for constructive, pluralist and critical discussion that has a role in contesting the opinions portrayed in news stories [26].

In this paper, we explore ways of making collaborative visual annotations to online news stories. In particular, we describe the design and evaluation of Newsr, a mobile app that allows readers to engage online in social and visual annotation of news stories with others. Newsr was designed to purposely avoid textual commentary and, instead, invites users to directly annotate, revise and even redact parts of news stories they find stimulating, contestable, debatable or problematic. We studied the use of Newsr across a four-week period with 15 participants. Each week, participants chose new stories to annotate from a larger pool of stories and then annotated these with their co-participants online. We interviewed participants during the study, and conducted post-trial focus groups to explore participant reaction and engagement with Newsr.

Our analysis of our engagement with participants, as well as the resulting annotated news stories, highlighted how people used Newsr to question the relevancy of certain types of news stories, challenge the authenticity of facts and sources, and used the written annotations to draw other participants’ attention to specific content. The findings of our study offer two contributions to the field of HCI in the form of design implications. First, we offer a novel method for engaging users with online news through collaborative annotation. Second, we highlight how Newsr challenged users to move out of their own filter bubble by presenting news in a new format that reduces selection bias.
RELATED WORK

We contextualise our work two-fold. First, through a discussion of news engagement practices that includes online news sites, sharing of news (including notions of "fake news") and critical engagement therewith, and secondly, against work in collaborative digital mark up.

Engagement with online news

The development of online news media has greatly impacted the way in which people consume and understand news. In the UK, the use of online news sites has tripled over the last decade, and as of 2017 it is estimated that 64% of the news is accessed online [32]. A key feature of many online news sites is support for commentary by readers. The most common form of engagement between readers in this context takes place in BTL comment feeds at the end of the articles, often attracting hundreds and, on the largest news sites, thousands of comments [10].

While comment feeds on new sites bring opportunities for engagement of readers around topical issues, it is also noted that BTL features come with their problems. Studies have highlighted the ways people use them to deliberately seek to invoke controversy [36] and engage in anti-social online behaviours such as trolling. As such it is not uncommon to see statements from users that might openly contest the rights of individuals discussed in a news story [30], to provide misleading information to make a specific point [31], and to continuously repeat the same argument despite being provided with evidence to the contrary [3]. Such engagements limit an otherwise more valuable public discussion space [16], especially around socio-politically sensitive topics. Consequently, comments might not accord with the messages journalists would like to get across [12]; or in the case of news stories that are biased these discussions might reinforce similar bias [4]. Aside from potentially being offensive and misinforming, ‘low quality’ commentary upon online news articles are additionally problematic in shaping other readers understanding of the content of the news article. As Anderson et al. [3] note, audiences to such ‘uncivil’ commentary may find the messages hostile and consequently make judgments based upon pre-existing values, dispositions, and perceptions as opposed to focusing on information presented; in doing so, they can reject being open to unfamiliar topics, positions and information [11].

A typical approach to resolving the challenges of ‘uncivil’ discussion in comment feeds has been to develop automated systems that facilitate moderation and remove ‘trolls’ [9, 35]. Studies have highlighted that users of online news sites might be more willing to engage in discussion if they are aware that a comment feed is moderated [38]. However, classifying the differences between a ‘troll’, a civil user, or a potentially community enriching social mischief maker [24] is highly challenging. Often formal definitions of ‘trolls’ are at odds with those of moderators, news editors and other users’ conceptions of what constitutes such behaviours [35]. Another approach to countering such behaviour has been “distributed moderation” systems, whereby many users participate in moderation of commentary [28]. While partly successful, such measures are still recognized as being problematic given the timeframe between posting, reporting and removal, allowing for conversation to occur [ibid]. Consequently, many online news sources now consider the removal of commentary sections at the cost of closing valuable public discussion and disempowering users in their legitimate pursuit for debate in the public sphere [18].

Online sharing of news and its consequences

In addition to the potential of online news commentary to derail critical conversation and generate misinformation, there is an increasingly recognised concern that ‘fake news’ and ‘echo chambers’ also dominate news consumption. As noted by Al-Rodhan [33] such ‘fake news’ is a hallmark of what might be considered an era of ‘post-truth’, defined as “relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief” [29]. Consequently, such post-truth politics and media hold “reliance on assertions that ‘feel true’ but have no basis in fact” [37] and often appeal to emotion and existing presumptions in motivating support (and consumption) over evidence-based or unbiased reporting. Notably such acts are likely to run free from verification or repercussions [33]. Such support and consumption of fake news is often bolstered through ‘filter-bubbles’ and echo chambers whereby individuals consume news by simply visiting their preferred news media source, ignoring alternate options in favour of content that follows their previously held beliefs [11]. Such echo chambers, it has been argued, also occur through reading those news stories favoured by their friends, whether this happens through background algorithms (e.g., through a news recommendation service or a Facebook news feed) or their own choice [27]. Consequently, differing information or opinion is not necessarily intentionally rejected but often unseen or unheard against an apparent vast amount of media supporting the contrary (and held) position.

Critical engagement with and social annotation of media

Moving beyond draconian options such as eliminating commentary, a growing area of research attempts to invoke criticality and novelty in how we consume and comment on online news and media. Such work is sometimes motivated by exploring ways for users to come into contact with content they may not normally do so. For example, on the back of a study of Twitter discourse surrounding socio-politically sensitive TV shows, Brooker et al. [7] called for social media design patterns that support alternate forms of news feeds around politically charged media; specifically, they call for interface techniques that enable users to see different layers of online discourse, and have them be confronted with views and opinions that might contrast with their own. Ideas such as this have been proposed as ways for users to break out of echo chambers on social media [15] or in public settings [14].
While the above examples illustrate approaches of confronting users with alternative views to one’s own, others have approached the issue of problematic commentaries through supporting more critical engagement with the content of the media itself. In the context of broadcast television programming, Feltwell et al. [13] fostered critical discussion around socio-political topics, such as welfare reform and poverty, through the design of ‘second screen’ applications. They present different applications that support active reading and annotation of TV shows, or social tagging of social media conversations that were occurring alongside such shows. They noted the ways in which engaging viewers in creating tags and annotations related to on-screen content supported reflection on the ways in which subjects were represented and interrogation of editorial choices. Research around user-generated tags and annotations also highlights their benefits for communication and navigation of content for other users [1]. Engaging users in the creation of tags also supports critical thinking around how they would be read and interpreted by others [ibid] while also engaging users in thinking carefully about the nature of the tagged content [5].

Research in collaborative digital annotation offers constructive ways to rethink user engagement with different media. For example, Hansen [19] looked at the central challenges in annotation systems that can bridge into the physical world. Carter et al. [8] investigated the annotation of public multimedia content through ‘digital graffiti’. A core aspect of this work was supporting handwritten annotations via PDA across three modalities: personal annotation for active reading; collaborative annotation that draws others to specific details; and social/public annotation that provides commentary for others. Heer et al. [21] present a web application to explore asynchronous collaboration in the context of information visualization. They report on mechanics that include view sharing, typed discussion, graphical annotation, and social navigation. Significantly, they describe the importance of “grounding” shared discussion to specific material on the same view, and how annotations are double-linked - from a comment to a view (and vice versa). Their analysis of graphical annotations explores how users create a pointing interface by scribbling on an “acetate layer” over the visualization. There are also commercial mark-up systems that are ready to use: Twiddla [23] provides an online whiteboard for classrooms with “real-time” markup and rich drawing tools (while retaining an accessible tool palette); Microsoft Whiteboard is a freeform digital canvas for creative collaboration between Windows devices with a similarly accessible tool palette. These can be contrasted with the online media site, Medium.com which allows users to respond with comments against the side of news stories anchored against the relevant text rather than BTL. This inspired the open source SideComments [2] that saw equivalent technology released for developers. Elsewhere, Snapchat allows personal annotation of published snaps through digital coloured pen and emojis. While these works do not address news media commentary directly, they offer collaborative ways to rethink user engagement with online news articles within a social context. Through the incorporation of shared annotation there is opportunity to explore ways to overcome the challenges faced by BTL feeds and encourage active and critical reading through ‘tagging’ media content. Furthermore, this prior work shows how we can draw others to, and focus attention to specific elements of news, rather than, for example, argue about peoples’ different positions.

**NEWSR**

We designed an app to explore how the social and visual annotation through graffiti-like interactions might engender new forms of critical engagement with the content of news stories. Based on the above literature, we were motivated to examine new methods of bringing readers into contact with unfamiliar news sources from outside their filter bubble, while also looking for new ways to annotate stories that would benefit from the immediacy of our devices, in that they are ever present allowing us to engage with news over time bit by bit, and benefit from their social connectivity. In doing so, Newsr also speaks to potentially lost forms of interactions with paper newspapers where they might be ‘read with your pen’ and annotated in that moment against the media itself (as in [8]). Following this, Newsr has four different modes (and stages) of interaction that are accessed from its menu bar.
(i) Selection Mode: Upon entering the application for the first time, selection mode presents users with newly available news stories that can be selected from a stack of stories. Users are presented with an image of just the main headline of the story, a header image (if there is one in the story) and the opening sentences of the story. We limited the presentation of the stories at this stage to encourage participants to make a choice on the story using its bare essentials. The stack has a swipeable card interface (see Figure 1a) intending the user to make an initial reaction on the news story and act on it. Here, swiping right chooses an article for annotation and graffitiing, whereas swiping left ignores the article. Once the user has gone through the selection process they are unable to revise their selections. The intention was to make the swiping a weightier decision and encourage the user to visit the gallery to view how ignored stories were graffitied by others.

(ii) Graffiti Wall: After swiping new stories, the Graffiti menu option (see Figure 1b) opens the graffiti wall which allows the users to review only the stories they indicated they had interest in by the swipeable interface. This wall keeps their chosen stories which can be visited like saved “favourites” as required. From here, users can choose a story and open it in the Graffiti Tool (see Figures 1c to 1e). At the end of every week, stories are ‘locked down’ so they can no longer be annotated (indicated by a graphic, as seen in lower row of Figure 1b).

(iii) Graffiti Tool: The graffiti tool is the main focus of engagement with the content of news stories in Newsr. The tool presents the user with different annotation tools inspired by both commercial systems (e.g. [23]) and longstanding practices of annotating and scribbling on physical newspapers in pen. From left to right the menu options are freehand graffiti, eraser, strikethrough, highlight, previous and next page (see Figures 1c and 1d).

Freehand allows users to draw using touch, over any part of the text or image, similar to the “acetate layer” [21]. This provides users with a way to circle, point to text or sketched writing. This interaction is extended by pinch-in and pinch-out standard zoom gestures letting users create graffiti at a zoomed-in scale (see Figure 1c).

Selecting either strikethrough or highlight and then touching words on the page (or alternatively dragging a finger across multiple words or sentences) applies these actions to the text (Figure 1d and 1e). Strikethrough and highlight also benefit from pre-parsing the stories using optical character recognition. This enables the reader to engage with any normally rendered typeface in an article, but also, text embedded in the actual images within the story. For example, a news story might use photos of a physical billboard (see Figure 1d), while those about social media often evidence the story using photos of tweets which can be annotated.

The eraser option allows the user to remove any graffiti that they have created, but does not work on others’ graffiti. As demonstrated as useful elsewhere [17] a users’ graffiti appeared in its own unique colour. This ensured that the user would not know who the graffiti belonged to, but that a certain colour belonged to one user. This was important since some participants had met at the recruitment company. Finally, the options for previous page and next page allow the user to step through the pages of a story and look at the graffiti present on any of the pages.

(iv) Gallery Mode: The Gallery menu option opens Newsr articles for viewing only. Options for editing the news stories (as present in the Graffiti Tool) are removed. Here, users can check graffiti on the stories that they previously ignored in the selection mode. The stories in this view are continually updated to show other users’ graffiti as it happens.

Selection and formatting of news stories

The online news stories in Newsr were specifically formatted for the app using trending articles, intending these to be relevant for research participants from the UK. Newsr sources included UK based tabloids, broadsheets and global news outlets. Newsr stories were selected around the theme of ‘othering’ where chosen stories were split roughly between those that reported on instances of othering, incited hatred towards specific populations (othering those populations) and/or, focused on more personal accounts of individual(s) who have been othered. These were collected by two of the researchers and maintained in a ‘living document’ with a further set of exclusion criteria applied to finalise the content. This included removing images that might be considered culturally insensitive.

To format articles for Newsr, online content was arranged into the front page and remaining pages organised around images and blocks of text. This resulted in a story with partially full pages with top and bottom margins providing a physical space for graffiti (e.g. see Figure 1d). This process was done by hand which involved a lot of “sizing up” by eye to gauge the right amount of space for each page. The result of placing content on pages in contrast to a scrolling feed ensured that the stories could be delivered on different mobile devices with identically positioned text, so the absolute location of graffiti was coherent across devices and therefore pointed to the same content. This helped avoid issues with different mobile phones e.g. as caused by font kerning variances.

STUDY DESIGN

We designed our study to explore how the design of Newsr engaged users with online news. The study was conducted in an authentic context over a four-week period in-the-wild. It included a diverse group of participants and selected topical and current news stories. The study involved participants attending an initial interview, using the Newsr mobile app in the interstitial moments of their everyday lives, taking part in a mid-point telephone interview and a final focus group (detailed below).
Initial interview and participant briefing: The initial interview was conducted in person at Northumbria University, UK. The first stage of the initial interview took approximately 15 to 20 minutes, and involved participants talking about the types of news outlets they frequently visited, the types of stories they were drawn to reading, and if they engaged in any existing commentary features of news sites, or if they shared or commented on news stories on social media sites. Following this, we spent a further 15 to 20 minutes setting participants’ devices up for the study. This included supplying them with login credentials for the app and demonstrating Newsr on their own personal mobile devices (which was a mixture of different Apple and Android smartphones). After this, we explained the structure of the study over the coming weeks, and provided participants with a reference guide that detailed the schedule of the study and how to use the Newsr app.

Weekly use of Newsr: Following the first meeting, we asked participants to follow a weekly routine of selecting, reading, annotating and viewing the annotations of other participants. For three days at the beginning of each week (Monday through Wednesday), participants were given three to four news stories at the start of each day. During these days, participants needed to choose any newly appearing stories they wanted to look at in more detail by swiping-right (to keep) or swiping-left (to ignore) the stories. Once they had done this, they could return to reading full news stories based on their selections and annotate these using the graffiti functions of Newsr. During the latter part of each week (Thursday and Friday) there were no new stories, and the only expectation was for them to do more reading and annotation. Depending on when they engaged with Newsr, they would also be able to view the annotations of other participants on the same news stories. Finally, at the end of each week (Saturday), that week’s stories were locked so that they could no longer be annotated in the graffiti tool.

Mid-point interviews: These interviews were conducted approximately halfway through the study (near the end of week two) and performed over the telephone. These were short interviews of approximately 15 minutes in length, where participants asked about their experience of using the Newsr app so far. We specifically asked questions related to the selection of news stories we had provided and how they came to choose certain stories for annotation. We followed this up with questions related to how the participants went about annotating the stories. These interviews also provided a point at which we could gauge whether participants were encountering any bugs or usability problems with the app.

End of study focus groups: At the conclusion of the study, participants were invited to take part in a focus group again held on university premises. The size of groups was kept purposely small (between 3 and 4 participants) to ensure all participants had a chance to share their experiences, but large enough to engender sharing of different experiences and opinions. Similar to the telephone interview, semi-structured discussion points were used to explore the experiences of using Newsr over the course of the study. However, greater emphasis was placed in the focus groups on discussing how participants engaged with other peoples’ graffiti and how, and if this changed their opinion of the stories. To help facilitate this discussion, we picked the pages of graffiti that the participants would find the most familiar. In concluding the focus group, we debriefed participants.

Participants
Fifteen participants took part in the study, all of whom were recruited via a local recruitment company. All participants attended the initial interview, telephone interview, and while one participant could not attend the focus group, they did not withdraw from the study. The recruitment criteria specified they were aged over 18 years, self-identified as a regular reader of online news, and own a smartphone produced in the last three years (for software compatibility). Participants did not know each other, with the exception of two participants who had met through the recruitment company.

The initial interview showed how there was a diverse range of reading interests. Participants differed in where they would typically receive their news from. These sources included The Sun, The Daily Mail, The Mirror, BBC News app, Sky News app, Independent and Guardian. Seven of our participants self-identified as female, and 8 as male, with an age range between 20 and 53. Participants were compensated for their time on the study at the rate of £12 per hour, which was calculated pro-rata based on the amount of time they participated in the study, up to £150 (payment to gain access to participants [20]).

Data collection and analysis
All of the interviews and focus groups were audio recorded, then transcribed for the purposes of data analysis. These were collated with all of the annotated stories from participants, along with analytics data from the web app related to the times when participants engaged with the app and which stories they selected at the start of the week.

Data analysis followed an inductive thematic approach as per [6]. We began our analysis while the study was in progress by introducing discussion points with participants around their more interesting annotations. In interviews, each participant’s graffiti allowed the research to question the choice of stories, and how they had interacted with each story. Correspondingly, the final focus group was structured around the most graffitied articles as well as those articles which had been edited by the most people in the group. At the conclusion of the study, images and transcribed data was analysed by three of the authors through a process of open coding. During this process, we sought to code connected data from different sources alongside one-another (e.g., annotations of specific stories alongside interview and focus group data where participants discussed these stories and annotations). Having coded the data, codes were compared, contrasted and combined as appropriate, from which six overarching themes were constructed. These themes

...
The swipe functionality of the app where participants could swipe-left to ignore an article and right to keep it, provided a means for participants to quickly sift through stories to put aside those they were most interested in. In the most part, participants commented on how this functionality was a useful way to have small glimpses of stories to engage with more deeply at a later point in time: “the ones I swiped right for, it was just that initial kind of headline or sort of picture format was... kind of captivated me to want to read further” [P14]; and “I think that’s nice to be able to just like, say, ‘No, I’m not interested in that’” [P11] and “I like the way you could swipe them and then come back later, which I did quite often if I was at work and I had five minutes I’d accept the stories and then I’d go back on the night time and actually do the comments” [P14].

Significantly, the small glimpses provided a source of intrigue. For example, in relation to a story about “illegal immigrants” loitering outside a retail store offering their labour for money, [P3, S4] explains, “I think I was just intrigued. I just wanted to know whether or not there was any proof to it. It seemed very specific that it was £50, but then I think when I did read it, I am sure it said something like it was just people hanging around outside of Wickes”. Here, intrigue was born out of some cynicism and a need to find out how a story happened. In another example, a front page exclaimed refugees were using air travel (see Figure 2b): “Yeah, because first of all I thought, so that means the cargo hold? I thought, because they can’t do it, it’s pressurised, and I thought, ‘How can they do that?’ so yeah, it was intriguing. And someone [else] wrote, ‘Intrigued’ at the bottom...” [P14, S22] (see Figure 2b) and similarly in a different group, “On this one I couldn’t quite believe how in this day and age people get... can sneak on to a holiday trip, I just can’t... I couldn’t fathom it” [P6, S22]. It was also apparent that looking at others’ graffiti formed part of the intrigue, for example, in discussing articles on gender inequality a participant commented, “those articles were particularly interesting because it’s a new emerging thing in the world really sort of transgender rights and gender neutrality in schools and in the widespread world, so I was quite interested to see what other people had to say” [P1, S17].

There was also further intrigue generated from recognising the source as being from a tabloid. e.g. “I was a bit intrigued because I don’t read tabloids, like over what was being said” [P5]. In these cases, this led participants to question whether the facts would back up the argument, “I wanted to see the way in which the newspaper itself decided to go with that story, whether it was just going to be a sort of benefits bashing typical Daily Mail style article, or whether there was going to be more in-depth analysis” [P1, S24].

Generally speaking, participants selected stories based on a more personal interest in the topic: “I’ve been quite interested in this stuff about sort of gender and sexuality” [P2] and “the gender pay gap was good for me, NHS Scotland because I’ve got a daughter who’s a nurse in Scotland and a nurse down here so that was really interesting” [P4]. However, the act of graffitiising such stories did change what participants chose later: “Now if something
comes up that I think, Oh, that’s quite similar to what I’ve read or commented on already. I’m not [swiping] it” [P3].

Questioning and Challenging Relevancy
Graffiti was created at different levels of scale in the news stories i.e. at the level of the whole article, pages, and at the smaller scale of paragraph, sentence, word, hyperlink, image, and individual letters e.g. for spotting spelling mistakes. At all these varying levels of scale, participants questioned and challenged relevance, both whether the story was newsworthy (see Figure 2a), and that the content of the article was appropriate and connected to the subject. Sometimes, participants used the whole page to indicate that a story was, in their opinion, irrelevant e.g. “NOT NEWS” [P12, S36] was scrawled across a front page over both text and photo. Margins also provided space for broad criticism of content: “Who cares” [P8, S13] (see Figure 2c) was placed in the margin below a photo from a story: ‘Burqas or bus seats?’ with another interviewee explaining, “Why is it newsworthy? [...] They could have chosen something much more relevant if they were looking at an issue like this, other than a photo of bus seats” [P5, S13].

The relevance of images was also challenged. In a story focusing on women’s health, participants noted photos of fashion models were overused, despite having little relevance to the health issue in question. This prompted one participant to write “Model again” [P4, S39] over a picture of a model (see Figure 2d), followed by “sick of pics now” over the top of another photo, as explained in interview “[I was] just sick of the pictures of this body that nobody has got apart from if you are a supermodel” [P5, S39].

As with the following themes, specific content in the articles was challenged through use of the highlight tool, and freeform drawing tool via circling of content, underscoring text and pointing with drawn lines and arrows. For example, in a tabloid article about a mother in receipt of state welfare (see Figure 2e), a participant explained “I highlighted things like ‘Ex-lap dancer’, is that really relevant?” [FG3, P3]. As another participant adds, “Yeah, I think I wrote the same thing, yeah. How is that relevant? I was just really angry reading this one…” and later explaining highlighting the words ‘midlands twang’: “Why is it relevant that she had a regional accent?” [P9, S24].

However, graffiti was not always disparaging, and occasionally, was used to praise content: “Sad Pic appropriate” [P4, S1] over the top of a picture of a wreath and “provocative image” [P3, S5] on a wrecked lorry which had been used in a terrorist attack.

Checking authenticity of facts and sources
Graffiti was used extensively by participants to challenge the authenticity of an article. For example, margins provided space to question whole pages: “Rubbish” [P3, S3], “hearsay?” [P11, S7], “all a bit vague” [S4, P3] and “robust study?” [P3, S39]. Specifically, the different mechanisms used to identify and link content to graffiti, enabled sources for claims in stories – such as hyperlinks to web sources – to be questioned. For example, “how reliable is this blog?” was used to highlight the use of a blog as a source [P3, S2] (see Figure 3a), as they explained, “Who owns England blog? I’m not sure how robust … I don’t know who England’s blog is, or what its credentials are”, and in circling the word blog with the comment “reliable?”, another participant explains, “I’ve commented similarly about how reliable is that blog that is being used, and whose figures are they?” [P5, S2]. Conversely, participants also used graffiti to challenge the complete absence of facts and sources in paragraphs, e.g. “proof!” [P5, S4] was used to challenge the highlighted words illegal immigrant. As P5 described, “Yes, so how did they know that they were migrants? It was just, I think, a very general story with no proof” (see Figure 3b).

Bringing attention to specific content
In bringing attention to specific bits of content, participants were able to express their own viewpoint. This graffiti often took the form of emotive language which was likewise associated with both text and photos. For example, the words “Tragic” / “Desperate” [P9, S1] (see Figure 3d) were used above and below a headline talking about the loss of life in a human trafficking tragedy. Later, “naive, so sad” [P11, S1] and “horrific” [P1, S4] (see Figure 3e) were used to further describe participants’ feelings about the events, with P11 explaining: “so sad that they know they’re going to go into this van, he didn’t take any water trusting that there was going to be some, [...] they’re so desperate that they need to leave [...] It just breaks my heart.”
Viewpoints were quite often concise, being indicated with short sentences or posed as rhetorical statements on pages, e.g. “whys” / “wheres” and questions e.g. “true?” were common. Occasionally, general agreement was as simple as drawing a tick against some content without accompanying words, while striking through complete words was a efficient way to indicate disagreement with text. This was important as participants highlighted an overall lack of space: “I was struggling to find somewhere to write, it was hard to find a space” [P14] and “not enough space, even though I got quite neat, eventually, after a bit of practice, but I still found it really difficult.” [P11]. This was especially important if you were annotating the story late in the day, as indicated: “if you were quite late to the story and four or five people had already written on it, it made it quite difficult to read” [P10].

Whatever the marks made by participants, being connected to others over the Internet allowed participants to disagree with other people’s graffiti, and in turn, the underlying content of the article. As described by a participant, “I’d seen somebody else wrote, ‘Agree’, under something else that somebody else had wrote [sic].” [P8] (see Figure 3c). However, these exchanges were limited, “I mean that’s pretty much the extent of a conversation you can have though, sadly there’s not enough space to allow you to go back and forth for very long, it’s usually just ‘Agree’, ‘Disagree’, ‘Rubbish’. And stuff like that.” [P1]. And some topics met with more agreement than others, such as racism, “Everyone agreed on that. Gender stereotype the same thing, I thought some of the things to do with sexuality and gender were a bit more divisive. I think you saw more opinions in that so you could agree and disagree more. Whereas again, stuff like racism and feminism is more agreed upon in our little group” [P1].

Participants were also keen to highlight specific language use. This was often used to mark where the choice of language was notable. Occasionally, participants used the highlighter to praise the choice of language e.g. the use of *Mexican labourer* was highlighted and linked to the graffiti “NOT illegal immigrant” (see Figure 3d). This participant went onto explain, “the story could have been a lot more negative but was actually quite kind of sympathetic towards the man who like tried to get into the country”, [P5, S1].

Participants used the highlight tool to identify a range of problematic words in the story about the mother wanting a horse. The words questioned by the participants were clearly “othering” the person at the centre of the story e.g. “I think I highlighted things like “Octomum” and the sort of terminology that you never really hear. The brumby benefits scrounger, like just the way it was written I was thinking it’s just The Sun, it’s like a tabloid… “, [P6, S24] (see Figure 2e, Figure 4a and Figure 4b).

**Wider reflection on the story**

Importantly, Newsr helped participants step back and reflect on the topics and wider issues in the news story. For example, participants related stories to other topical news they had read outside the app: having written “Shocking compared to other BBC wages” (see Figure 4c), the P14 explained, “I’d seen something on BBC news recently about… it was listing the top salary earners like on 100,000, 200,000, 300,000 and you’re like, how the hell can you justify that with somebody who’s been doing it for a long time and is earning £16,000?” [P14, S9]. They also related news to wider societal problems, for example, in a story about rehousing victims of a fire, one participant wrote, “London just for the rich” [P9, S2], and later explained, “I commented on this Grenfell Tower thing, almost 2,000 homes near Grenfell Tower, lying empty as the authorities struggled to rehouse victims. […] while it is awful, I think it’s an ongoing housing issue in London boroughs. It’s not an easy, or a quick thing to tackle”. Additionally, expressing these views often led to participants thinking more actionably about the topics, and calls for action became part of this wider reflection. In responding to people needing a home after the Grenfell fire, a participant wrote “Faster action needed” [P14, S2] (see Figure 4d), while responding to an article on modern slavery another writes, “Yes take action” [P0, S26].

**Judging people in stories**

The graffiti created by participants was often carefully considered both in respect to what was said, and also where it was placed. However, it is important to acknowledge that...
Newsr could be used to make personal judgements. The few instances where this happened were made in articles that centred on an individual. This involved some name calling and mockery. For example, on the front page of the story about the mother wanting a horse [S24], a participant had written, “that’s a funny looking horse” [P1, S24] and “why the long face”. The editors had also reused the same photo on a subsequent page which was then given the remark “stupid” [P5, S24] with arrows pointing to the person. Correspondingly, the participants initial reaction was to be incensed at the person in the article e.g. “Not what £ is for” [P4, S24], while another wrote, “Get a job first stop claiming” and “Should lose all benefits” [P14, S24], explaining later, “yeah, I commented saying, ‘Should lose all benefits’, I was fuming, I was really angry, I was like, how dare they spend the money we’re providing to buy a bloody horse when it’s supposed to be there for living”. However, in the telephone interviews it was clear that participants were more contrite; with the first of those mentioned pointing out their graffiti was fairer, since it indicated that £26,000 is “not a lot for 8 kids”. However, celebrities featured in the stories also received kinder graffiti: one article claimed that Dame Helen Mirren while fronting an ad campaign for L’Oreal, said their moisturiser ‘probably does f--- all’. This elicited reaction including “Love THIS!” [P12, S5], while photos that showed a close-up of her skin were given a “wow” [P4, S5], and on a similar photo, “something is working looks fab” [P11, S5] (see Figure 4e).

**DISCUSSION AND IMPLICATIONS**

Our work has attempted to rethink how users select, read, reflect and comment upon online news articles. This was achieved through an alternative approach to news article selection, providing news articles from multiple sources and through the addition of digital graffiti styled annotation. This allowed participants to view others’ ‘active reading’ and encouraged their own through drawing attention to specific details, as opposed to across the broader article. In doing so we avoided commonplace BTL comments and facilitated more critical engagement with online news media.

Our findings show how the design of Newsr has value in this design space. First, we encouraged participants to look over the front pages of online news stories from sources they might not normally consider, and upon topics that did not always align with the position of their regular news source. In contrast to our concerns that participants would likely identify and select stories from newspapers they were familiar with and read, participants were interested in the front pages. They made selections, partly out of intrigue, because they wanted to see how stories developed, what had happened, whether it was supported by facts, and to see how others in their group had responded. Second, participants engaged with the news stories in new ways. Participants created graffiti to critique different scales of content i.e. from pages, to individual letters, motivated by different purposes. Annotations were made in margins to critique the relevance of the article, page or photo, or addressed minutia of content. This included showing how language was inappropriate, and even highlighting typos. Importantly, annotations shaped the discourse surrounding a story, as participants responded to each other’s annotations, ultimately choosing different articles because of their experiences.

In the following sections, we synthesise the learning from our findings across a series of challenges our participants faced in experiencing Newsr; while we pose these as challenges, we note that each one provided productive frictions [25] that promoted engagement with unfamiliar stories, engagement with the claims of news, and engendered both absorbing of and contributions of annotations.

**Being challenged to choose new news stories**

Newsr challenged participants as they picked news stories encouraging them to expand their filter bubble [34]. It did this by altering the presentation of news stories through removal of source information, while still providing the user agency to choose multiple stories from a larger selection. This formatting was important and in selecting content around “othering” which often use divisive imagery and phrases, our stories no doubt “court” response. As such, it is difficult to know what interest ordinary news stories might receive. We can therefore imagine new systems where the source might be better masked than ours. e.g. using different fonts, colours, CAPS and punctuation marks in headlines, or even language processing (as in [9]) will increase homogeneity in the front pages. This might circumvent selection bias rather than allow users to vet content. In our study, participants were able to guess the source paper indicating that our formatting did not mask the origin well enough. However, despite knowing the origin (or at least thinking they did) participants selected stories outside their own filter bubble for a challenge. This also poses the question about how much should be revealed. In-keeping with our stack of news stories and swiping mechanic, future recommendation systems might also vary the stories (as is commonplace) but also what is revealed based on the users’ previous selections, reducing potential “echo-chambers”.

**Being challenged and challenging the legitimacy of news**

The participants in the study appropriated Newsr finding their own methods to question the legitimacy of the news. In Questioning and Challenging Relevancy participants decided whether the topic was newsworthy and remained on track for the rest of the writing. In Checking authenticity of facts and sources participants were able identify and highlight the lack of sources and validate existing sources. While there are existing systems [9] which can be used to fact check these could be integrated into systems like Newsr. However, these systems will benefit from being linked with the mechanics used by readers to bring attention to content i.e. the highlight, drawing circles and using arrows. These mechanics might signpost potentially problematic parts of the story for others, while simultaneously informing users of additional knowledge e.g. highlighted facts might be checked automatically and automatically added to invite critique. Alternatively, we might imagine users can further
“tag” content (like those in [8]) - adding sources of information to both support and contradict content.

**Being challenged by the annotations left by others**
The annotations left by others were concise, taking the form of single words, short sentences, and even rhetorical questions. These were hand drawn and consequently hard to read, and with the exception of a few instances where neat writing had been created with a stylus, the graffiti was hard to read. Consequently, the graffiti left by users was difficult to understand – both in terms of reading what was written, but also in terms of understanding what was written. For example, if a participant had highlighted a selection of words without explanation, the motivation behind the graffiti might be ambiguous. The findings describe the many reasons for drawing attention to content, from circling to agree with something, to calling out problematic language. However, the ambiguity of these markings need not be a bad thing; indeed, such ambiguity can be a useful resource for design [12]. In our example, the value is in drawing attention to significant article elements. This plays a part in helping people with opposing views find a “middle ground” as opposing positions can note an element as contentious, yet refrain from explicitly providing opinion. All these above acts of reading and annotation contribute to how users reflect on content, such as helping participants relate stories to wider societal problems, and other topical news stories, which became another source for alternative views.

**Being challenged to contribute your own annotations**
The findings showed how participants created graffiti against different types of content e.g. photos, paragraph, sentences and words. Importantly, this was added piecemeal, with most users annotating bits of the story as they read it and a page at a time. As such graffiti, sometimes evolved across the article as readers changed their opinion. This also provided a unique challenge for participants, as they needed to find a space to distil their point. Thus, whoever accessed content first was often the one to occupy the space with thoughts. Instead of offering typed text entry that would use space concisely, we instead had the freeform graffiti tool which helped elicit different forms of responses, such as those pointing mechanisms seen [21]. As such the annotations are unlike existing comment systems for news and limit the ability for conversation and reduce opportunity for arguments. Instead, readers restricted themselves to simple annotations like a tick, or wrote agree/disagree to indicate their thoughts. In thinking about graffiti in the article, we must choose carefully where we provide space for annotation and formatting is a critical step here. As in other parts of these systems compromise is important, such as, providing just enough space for critique vs. too much space. It is important to note the formatting was done by hand and a “best guess”. We can see opportunities for algorithms to perform this role better since some stories were given more room than others which invited more verbose annotation and might afford arguments.

Finally, the findings describe instances where Newsr has been used to make judgments on individuals. However, there were many instances where the reader showed more empathy. On reflection, we might consider that some articles are better suited for eliciting more constructive debate.

**CONCLUSIONS**
We report on the design of a mobile app, Newsr, which encourages readers to interact critically with news stories drawn from both familiar and new sources of news. We conducted an evaluation of the app in-the-wild over a period of 4 weeks drawing upon current news stories. Our findings are based upon qualitative analysis of data gathered from interviews and focus groups with 15 users. Our work contributes to knowledge on critical engagement with news looking primarily at how users can annotate news stories in new ways. This also contributes to work on the social annotation of news, work that explores and challenges filter bubbles. Our findings provide a set of implications for the design of news reading applications, that can be applied to designs and research that addresses how readers think about news sources, and in work that explores new ways to interact with the news. These design implications are presented in terms of challenges in such interfaces, challenges we argue, encourage critical reflection.

We have already discussed how the curation of stories will affect user responses, and acknowledge our process selected stories around ‘othering’. As has been demonstrated, good content selection for presentation to users is key to engagement and reflection, and considering existing journalistic practice this is somewhat unsurprising as all news sources are edited and curated to evoke reactions from readers. Therefore, this is a limitation of our study, and future work in this area may wish to explore content selection from a more diverse range of stories.

In this paper we have looked at two modalities: the swipeable interface for choosing stories, and annotation through graffiti. We believe these are complementary. Together, they push readers to explore new sources as the graffiti itself becomes a draw to the article. The swiping interaction is quick, and together with the piecemeal annotation of news (both in the type of content and how content is annotated), allow annotation over time, making Newsr a better fit with reading with your pen and reading on the move.

**ACKNOWLEDGEMENTS**
This work was supported by the RCUK grant ES/M003574/1 “CuRatOR: Challenging online feaR And OtheRing”. CuRatOR is funded through the Empathy and Trust in Online Communicating (EMoTICON) funding call administered by the Economic and Social Research Council in conjunction with the RCUK Connected Communities, Digital Economy and Partnership for Conflict, Crime and Security themes, and supported by the Defence Science and Technology Laboratory (Dstl) and the Centre for the Protection of National Infrastructure (CPNI). Data underpinning this paper is available from the authors.
REFERENCES


