Marshall, Joe and Linehan, Conor and Spence, Jocelyn and Rennick-Egglestone, Stefan (2017) Throwaway citation of prior work creates risk of bad HCI research. In: CHI 2017: ACM CHI Conference on Human Factors in Computing Systems, 6-11 May 2017, Denver, Colorado, USA.

Throwaway Citation of Prior Work Creates Risk of Bad HCI Research

Joe Marshall

Mixed Reality Lab

School of Computer Science

University of Nottingham

Nottingham, UK

joe.marshall@nottingham.ac.uk

Conor Linehan

School of Applied Psychology
University College Cork,
Cork, Ireland
conor.linehan@ucc.ie

Jocelyn Spence

Mixed Reality Lab

School of Computer Science

University of Nottingham

Nottingham, UK

jocelyn.spence@nottingham.ac.uk

Stefan Rennick Egglestone

University of Nottingham

Nottingham, UK

stefan.rennick_egglestone@nottingham.ac.uk

Abstract

In CHI papers, citation of previous work is typically a shallow, throwaway action that demonstrates little critical engagement with the work cited. We present a citation context analysis of over 3000 citations from 69 papers at CHI2016, which demonstrates that only 4.8% of papers cited are presented as anything other than uncontested fact. In 43% of CHI papers sampled, we found no evidence of any critical engagement. Lack of discussion and critique of previous work can encourage the spread of misunderstandings and errors. Authors, reviewers and publication venues must all change practices to respond to this failure of scholarship.

Author Keywords

Bad HCI; Citation context analysis; referencing

ACM Classification Keywords

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

Introduction

This paper presents the argument that the way in which researchers talk about previous research in CHI is problematic. CHI papers often present the findings of studies that they refer to as simple facts. Sometimes study details are described, such as methods, findings

and implications. Rarely are methods, findings or implications questioned, critiqued or analysed in detail in CHI papers. There are many reasons why this culture may have emerged. For example, the interdisciplinary nature of CHI requires researchers to read broadly, and space restrictions on conference papers can encourage authors to quickly get to the empirical work. We argue that, regardless of motivations, lack of critical practice in reviewing previous work has, and will continue to, undermine the quality of research presented in CHI.

Of course, this argument may be criticized as conjecture in the absence of a rigorous empirical analysis of writing practices of CHI authors. Therefore, we carried out a citation context analysis of over 3000 citations from 69 papers from CHI2016, chosen to cover the breadth of topics and subdisciplines represented in the programme of this latest CHI conference. Findings suggest that only 4.8% of texts cited were critiqued, analysed or questioned, and that few papers included any analysis of prior work.

We argue that the presence of a culture in which previous work is understood as uncontested fact is dangerous, as it can allow misunderstandings, misrepresentations and simplifications to propagate as "facts" that are then built upon by subsequent work. In particular, we believe that this is an underlying cause of multiple failures of interdisciplinary working in HCI which we detail in accompanying publications [8,9].

Background
In questioning whether the writing practices of CHI researchers are unusual or problematic, it is necessary to understand first the writing practices of researchers and scholars more broadly. It is necessary to consider the function of the literature review within an academic paper, and the reasons why we cite specific papers within those literature reviews.

Why do we write literature reviews?
This might seem like a very basic question. All scholarly communication and dissemination contains some form of literature review, background, or introduction section. Literature reviews may fulfill many different functions in academic papers and may vary across disciplines [1]. The norms and conventions around reviewing and citing previous work are rarely formalized, codified or strictly enforced in a discipline. In questioning the function or purpose of the literature review in a CHI paper, it is worth examining the Guide to a Successful Paper or Note Submission [13] published on the CHI2017 website.

"To demonstrate the originality of your contribution you should make sure to **cite prior work** *(including your own) in the relevant area. If possible,* **explain the limitations in this work** *that your contribution has overcome. Make sure also to cite publications that have had a major influence on your own work. Lack of references to prior work is a frequent cause for complaint – and low rating – by reviewers. At the same time,* **long lists of reference does not show engagement with previous scholarship***."* Similarly, many scientific disciplines that publish research in brief papers often advocate the use of a short three-paragraph structure with very specific functions for each paragraph:

"The first paragraph should be a very short summary of the current knowledge of your research area. This should lead directly into the second paragraph that

Paying homage to pioneers
Giving credit for related work
Identifying methodology, equipment, etc.
Providing background reading
Correcting one's own work
Correcting the work of others
Criticizing previous work
Substantiating claims
Alerting to forthcoming work
Providing leads to poorly disseminated, poorly indexed, or uncited work
Authenticating data and classes of fact (physical constants, etc.)
Identifying original publications in which an idea or concept was discussed
Identifying original publication or other work describing an eponymic concept or term
Disclaiming work or ideas of others (negative claims)
Disputing priority claims of others (negative homage)

Table 1. Garfield's reasons for academic citation [5]**.**

summarises what other people have done in this field, what limitations have been encountered with work to date, and what questions still need to be answered. This, in turn, will lead to the last paragraph, which should clearly state what you did and why." [11].

Thus, it appears that the function of a literature review is generally to demonstrate originality, to demonstrate improvement upon previous work, and to convince the audience that the research is valid and worthwhile.

Why do we cite specific papers?
Bornmann and Daniel [1] suggest there are two contrasting theories used to explain citation behavior:

The "normative" theory, suggests that scientists cite papers in order to acknowledge the influence of the work of colleagues. In this view, a citation represents a signal that the cited work has had intellectual or cognitive influence; it points the reader to work they may not have encountered before, some of which may hold further interest for us; and it provides peer recognition of the place in which the idea originated, as a sort of admission of intellectual property.

In contrast, the "constructivist" theory of citation behavior suggests that intellectual content of articles has little influence on how they are cited. In this view, the scientist is an actor whose role is to persuade the academic community of the truth and importance of their work. From this perspective, citation is a persuasive tool used to demonstrate how new work is an advance on previous research.

In reality, these contrasting views are actually complementary, and simply describe two valid

categories of reasons we have for citing previous work [5]. Indeed, Table 1 shows a comprehensive list of potential reasons for citation (from [5]). The important lesson in the context of the current paper, is that we expect to see a variety of types of citation in a paper, signifying different ways that we engage with previous work, from the normative to the rhetorical.

Citation context analyses
Citation context analysis is a research method that allows for examination of the relationship between cited and citing papers. Procedurally, it requires researchers to manually code the text around a citation, according to a set coding scheme. Bornmann and Daniel [1] present a review of 30 citation context analysis studies in a wide range of different fields. The review provides empirical evidence of the citing behavior across many scientific disciplines, providing valuable context for our study of citation behavior in HCI. There is a caveat in interpreting those studies, which were typically undertaken for the purpose of understanding whether a citation is a valid measure of academic quality or influence, a separate question from that posed in the current paper. Nonetheless, the findings of that study provide the only relevant data we could find with which to compare behavior observed in CHI2016 papers.

Analyses of many disciplines, particularly in the physical sciences, concluded that citation behavior was largely normative. In other words, citations were most commonly made to papers that were relevant to, and influential upon, the citing paper. However, there are a number of other reasons for citing, summarized by Bornmann and Daniel [1]. We present a discussion of their conclusions under three headings (cursory, descriptive, critical), with the intention of describing the

way in which citations are used in a text. Note that these headings are not mutually exclusive and have been grouped into the three categories by us rather than by the original authors to more directly answer the research question raised in the current paper.

CURSORY

A surprisingly large proportion of citations in all studies (representing a range of 10 percent to 50 percent of citations in the studies reviewed) could be described as *perfunctory*. This category describes citations that mention work without additional comment, make redundant reference to cited work, or mention work not strictly relevant to the citing paper. A similarly large proportion (range: 5-50%) could be described as *assumptive citations*. This category describes citations that refer to assumed knowledge that represents general/specific background, refers to assumed knowledge in an historical account, or acknowledges pioneers. Citations of the *persuasive* type (range: 5-40%) describes citations made in a "ceremonial fashion" or where the cited work is authored by a recognized authority in the field.

DESCRIPTIVE

Citations labelled by Bornmann and Daniel as *conceptual* (range: 1 to 50% of citations in the studies reviewed) fit within our descriptive category because they refer to the presentation of definitions, concepts, or theories borrowed directly from the cited work. *Methodological* citations (range: 5-45%) refer to situations where the citing author identifies the use of materials, equipment, practical techniques, tools, analysis methods, procedures, or design directly copied from the cited work.

CRITIQUE

Citations labelled by Bornmann and Daniel as *affirmational* (range: 10-90%) describes the citation of work in a positive manner, but in more detail than a simple mention or description. Examples include where the citing work confirms the findings of cited work, where the findings of citing work are supported by cited work, where the contribution of the citing work depends centrally on the cited work, or where the citing work is strongly influenced by the cited work. Citations of the *contrastive* type (5-40%) describes a citation made in order to present a contrast or alternative between the citing work and the cited work, or to contrast other works with each other. Citations of the *negational* type (1-15%) describe situations where the citing work disputes some aspects of cited work, the citing work corrects or questions the cited work, or the citing work negatively evaluates cited work.

In analyzing citation practices in HCI, we should expect a spread of all these citation types. Studies reviewed by Bornmann and Daniel found that there were relatively more citations of the cursory than critique types across all disciplines. However, disciplines such as physics report quite high percentages of citations that could be described as in some way critiquing the cited work [1].

How does CHI Cite Prior Work?

We argue above that CHI has a tradition of 'throwaway', shallow citation of prior work. In this section we present evidence for this strong assertion in the form of results of a study carried out to establish how CHI papers refer to previous work.

Methodology

SAMPLE

To obtain a sample of papers representative of the full breadth of topics and paper types at CHI 2016, we downloaded the first paper from half of the thematic sessions at the conference. This gave us a sample of 3,183 citations covering 69 papers (13% of all papers).

ANALYSIS PROCEDURE

Two researchers independently read all papers in the sample. Since the focus of the current paper is on understanding practices in reviewing literature, we confined our analysis to the introduction section, plus 'background', 'context' or 'literature review' sections, which in all papers sampled, a) followed directly on from the introduction, and b)contained the vast majority of citations. Each in-text citation was labelled with one of our five pre-determined codes. It should be noted that there are many different ways that we could have coded these data. Indeed, Bornmann & Daniel [1] identify that thirty citation context analysis studies across multiple disciplines each used a different coding scheme, determined by the authors of those studies. Codes used in our study represent a version of the category headings we describe in our discussion of Bornmann & Daniel above (cursory, descriptive, critique). We split the cursory category into two ("list" and "work exists") because we noted in initial reading that "work exists" citations were often lists of multiple papers. Similarly, we split descriptive into "supports a fact" and "described", noting that citations are often of the form "the sky is blue [reference]", which provides a minimal idea of results of the cited paper but a limited description of the study that is the basis of that result.

The codes each provide a simple description of how the cited text was discussed by authors of the paper in which the citation was made. Each code is listed, with overall category of that code (cursory, descriptive, critique) in brackets after. Codes used were:

List (**cursory**) – work is cited in a list, with no further comment or detail on the individual text.

Work exists (cursory) – the citation is an example that work exists on this particular topic, with no further discussion. It is mentioned individually, not only in a list of other papers.

Supports a fact (descriptive) –cited to justify a factual statement made. No detail or discussion is presented on research from which the fact is derived.

Described (descriptive) – Work cited is described, including any of its justifications, methods and findings. The research is presented as valid and reliable and no questions, comments or critique are advanced.

Analysis / critique (critique) – the work reported in the cited paper, including any part of its justifications, methods and findings, is affirmed, contrasted, or contested. As described above, this does not mean that the author is presenting a negative view of cited work, just that they in some way engage or comment on the work cited in a way that acknowledges it as something other than an uncontested fact.

We have made no comment on whether work cited was relevant. This is purely an analysis of how previous work is discussed in CHI2016 papers. All data is provided in supplementary materials.

	Number of citations	Percentage of total (%)	Cumulative percentage (%)
In list	638	21	21
Work exists	268	9	29
Supports a fact	849	28	57
Described	1179	38	95
Critiqued	146	5	100

Table 2 Percentage of citations of each type.

Results

Reviewers found 3,183 citations. Discarding 103 citations which were not to academic work 3,080 citations were classified. Cohen's κ was run to determine if there was agreement between the raters, which gave moderate agreement κ = .423 (95% CI, .401 to .445), p < .0005. We note however that one rater was clearly more lenient as to what they considered to be meaningful critique. Due to this, we decided to consider as critique the most generous possibility, that any citation which either rater marked as being critique was considered to be a critical citation. We further combined the rest of the results along the same generous lines, considering each citation as being in the highest category using the ordering "in list", "work exists", "fact", "described", "critiqued". We also considered distribution of critique citations between papers to see whether critique citations were concentrated in certain papers. See Table 2, Table 3 & Figure 1 for results. The key findings were:

- Only 4.74% of citations presented critique or analysis of previous work
- 95% of citations are presented as uncontested fact.
- 57% of the citations in our sample do not even discuss method or results of studies cited.
- A majority (64%) of papers contained one or fewer citations classified as 'critique'.

- 'Critique' is highly concentrated: 52% of critique citations are in just 12% of papers.
- In the terms used by Garfield[5], citations at CHI2016 were overwhelmingly used to pay homage and confer authority on cited works, as well as pointing to further reading, rather than criticizing, critiquing, substantiating, disputing or correcting.

Figure 1 Critique is concentrated in a small number of papers

Critique citations in paper	0	1	2	3	4	5	6	7	8	10	12	15
Number of papers	29	14	5	5	5	1	3	1	1	1	2	1
Percentage of all papers	43%	21%	7%	7%	7%	1%	4%	1%	1%	1%	3%	1%
Percentage of critique citations	0%	10%	7%	10%	14%	3%	12%	5%	5%	7%	16%	10%

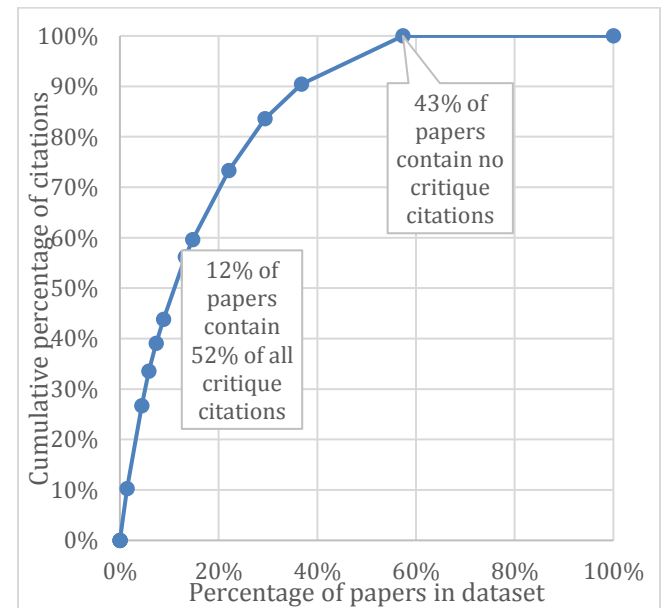Table 3. Distribution of critique citations across papers

Previous citation context analyses have found huge variation across disciplines in the proportion of studies that can be classified into our cursory, descriptive and critical categories. However, the proportion of citations in our study that are cursory appears to be unusually high, and the proportion that are critique is surprisingly low. Taking the lower bound described by Bornmann & Daniel [1] of citations which fit our critique definition (affirmational, contrastive, negational), we would expect approximately 15% critique citations. What we find from this analysis is that at least in writing papers, the tradition of CHI is very much one of citing prior work as fact, with minimal context or analysis.

Why is this a problem?

We believe that failure to understand and discuss prior work is already leading directly to poor quality research. A particular risk is the citation of work from other fields, where CHI's tradition of citing as fact comes into conflict with complex and not easily summarised ideas from other disciplines. As an interdisciplinary team covering specialisms from Psychology to Performance Studies & Computer Science, we were each able to identify such failures in our areas of expertise:

1) In [8], we describe how exertion gaming work mis-cites a single study of obesity and video gaming. The cited study *does not find a linear correlation between gaming and obesity*. Beginning with a massively simplified 'supports a fact' style citation in 2007, this paper is subsequently cited in HCI multiple times, with each citation distorting the fact further, until in 2009, it is cited to support the 'fact' that *videogames cause obesity*. In the context of

work aiming to alter videogames to 'cure obesity', this fundamental error means that the work cannot succeed in its stated design goals.

2) Concepts from performance studies are widely misused in HCI work, which leads to lack of clarity in terms of what the HCI work is actually referring to by words such as 'performance' and 'performativity'. They are also widely conflated with the 'performance' metaphor used in the work of social scientist Erving Goffman [12].

3) Work on computerized therapy which presents 'Cognitive Behavioral Therapy' (CBT) as the only or best way to do therapy. This misunderstands the therapy literature which increasingly supports a hypothesis that the 'named approach' to therapy is not particularly important in comparison to differences such as therapist personality [2]. It also neglects to consider current research into computerized versions of CBT (cCBT), which suggest that computerized CBT is not in any way the same as therapist led CBT and does not have similar efficacy [6]. We originally found this error in influential and heavily cited HCI work on cCBT, in which it is presented as fact; it is repeated in papers which cite that work.

4) Affective computing work which states that '93% of all communication is non-verbal' as fact. This popular myth comes from a misunderstanding of work by Albert Mehrabian [10] which studied situations where single words with positive and negative valence were said whilst giving positive or negative facial expressions with the opposite valence to the words. In this situation, if someone says a

positive word (e.g. 'lovely'), with negative facial expression or tone of voice, participants in 93% of cases saw this as a negative communication overall. This myth continues to be presented in published HCI work as fact, for example "*nonverbal behaviours, such as gestures, facial expressions or the way we use our voice, play a more significant role during an interaction than its verbal counterpart*" [4], citing either Mehrabian, or affective computing work which can be traced back to Mehrabian.

We present these 4 case studies in detail in an accompanying paper [9]. There are surely further examples in areas of work we are less expert on.

It is possible to argue that our focus on critique (and the lack of it) is inherently based in a 'scientific' model of research, i.e. that we are wrong to argue that good work must challenge, analyse or falsify prior work, and that, for example, practice based design work can still produce quality design whether or not it is founded in strong understandings of cited research. However, we argue that much work in CHI makes strong normative claims as to the goal of designs being demonstrated. In such work, arguments made in introduction and review sections of papers are key to demonstrating the potential of the research to successfully attain such goals and to succeed on its own stated terms. If motivation of work and alignment towards stated goals misunderstands or mis-states prior research, this can lead to pointless design, which inevitably cannot achieve its stated design goals. As such, we believe that review of prior work should be accurate and in depth. Essentially, we believe that irrespective of one's model of research, it is good academic practice to read

in-depth the sources which we are citing, and demonstrate in some way in our resulting work that we have read these sources.

How Can We Inject Critique into CHI?

We believe that CHI is sorely lacking in critical engagement with literature. We believe that to fix this, three key things need to occur:

- Critique of HCI research must happen.
- Critique needs to be ongoing, both during writing, in review process and after publication.
- Published critique must be situated at the core of HCI, not hidden in a critique paper ghetto.

In this section, we suggest four ways in which HCI writers, reviewers and publication venues could change to mitigate the problems described above:

As readers, we should be critical about cited 'facts'
The underlying issue with the obesity and videogame citation problem described above is that authors were able to present as 'fact' claims that were not supported by the evidence cited. This was made worse by the simplified nature of the facts presented, which led to further authors making more distorted claims. As readers, where we see facts presented with little detail, we must read source materials in order to evaluate such facts, and understand their limitations and nuances.

As writers, we should describe the work we cite
As Cozzens [3] and Bornmann and Daniel [1] suggest, citations have many purposes, both in persuading people of the quality and sound basis of an argument, and in performing other roles relating to the wider nature of academia as a social system. We believe by

describing in detail key work that we cite, and particularly by being clear about the assumptions and limitations of that research, readers will be less likely to be led into false beliefs about the findings of that cited research and to propagate them in their work. As a further benefit, this is harder to do without reading the source article in depth, and would perhaps have a role in helping avoid misconceptions in the first place.

As publishers we should improve citation clarity
CHI and many other HCI venues currently use ACM style numbered referencing. At major conferences, authors are incentivized to use unlimited numbers of such references. For example, in our dataset, one paper cites as a single group of citations: "*1, 5, 7, 13, 47, 74, 78, 79, 97, 102, 104*" [7]. Given the prevailing PDF format used for papers, it is a laborious manual cross referencing process to look up each citation. Even if the reader has a good knowledge of the related literature, they are unlikely to look up all 11 citations in order to understand which papers are being cited and whether they are being correctly represented. There are multiple ways to fix this situation, the simplest being the use of Harvard style citation, where it is relatively clear to readers who is being cited. This move would also discourage excessive quantities of citation being used to simply say that some work exists in a field. For example the above citation would be "(Allhutter 2012, Bardzell 2010, Bardzell & Churchill 2011, Bath 2014b, Draude et al. 2014, Light 2011, Lucht 2014, Marsden & Kempf 2014, Paulitz & Prietl 2014, Rode 2011, Rommes 2014)". Alternatively, cross referencing citations as links or comments in the PDF file as some publishers do could mitigate the issue for online readers.

As reviewers, we should critique related work sections
The CHI "Guide to a successful paper or note submission" [13] states that "*To demonstrate the originality of your contribution you should make sure to* **cite prior work** *(including your own) in the relevant area*". They even directly encourage critical engagement with prior work: "*If possible, explain the limitations in this work that your contribution has overcome. Make sure also to cite publications that have had a major influence on your own work.*" In our experience, reviewers often pick up on missing prior work in reviews, but it is very rare to see any major criticism of the quality of the analysis of cited prior work beyond comprehensiveness, or discussion of whether citations are appropriate or correct. Reviewer instructions should make it clear that reviewers must follow up citations that they are uncertain about and read source material. Further to this, reviewers should specifically consider whether the motivation of papers is well founded, to avoid situations where people do work which is based on objectively false assumptions and, which hence serves no useful purpose (see [8]).

Conclusions
Our analysis of a large sample of citations from CHI 2016 shows that 57% of citations did not describe any details of research cited. Typically, a statement was made, accompanied only by a citation, and there seems to be no expectation that this will be followed up or questioned by the reviewer, the reader, or future researchers. This behaviour gives statements in HCI papers an over-authoritative role. We believe that the prevailing style of citation in CHI has led to a situation in which cited work is misrepresented, oversimplified or exaggerated. Worse, over time such misrepresentations can have serious effects on the general direction of

research areas, leading other researchers into dead ends of well-intentioned but essentially fruitless work. We need to fight this both by embracing critical engagement with prior work in our own writing and by actually starting to consider prior work sections of papers and articles as a significant part of a publication that should be subject to the same evaluation as the rest of the presented research work.

References

1. Bornmann, L. and Daniel, H. *What do citation counts measure? A review of studies on citing behavior*. 2008.

2. Budd, R. and Hughes, I. The Dodo Bird Verdict-- controversial, inevitable and important: a commentary on 30 years of meta-analyses. *Clinical psychology & psychotherapy 16*, 6, 510–22.

3. Cozzens, S.E. What do citations count? the rhetoric-first model. *Scientometrics 15*, 5–6 (1989), 437–447.

4. Damian, I., Tan, C.S. (Sean), Baur, T., Schöning, J., Luyten, K., and André, E. Augmenting Social Interactions. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, ACM Press (2015), 565–574.

5. Garfield, E. When to Cite. *Library Quarterly 66*, 4 (1996), 449–458.

6. Gilbody, S., Littlewood, E., Hewitt, C., et al. Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. *BMJ (Clinical research ed.) 351*, 5 (2015), h5627.

7. Marsden, N. and Haag, M. Stereotypes and Politics: Reflections on Personas. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, (2016), 4017–4031.

8. Marshall, J. and Linehan, C. Misrepresentation of Health Research in Exertion Games Literature. *CHI '17 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (in press)*, (2017).

9. Marshall, J., Linehan, C., Spence, J., and Rennick Egglestone, S. A Little Respect: Four Case Studies of HCI's Disregard for Other Disciplines. *In submission, alt.chi 2017*, (2017).

10. Mehrabian, a and Wiener, M. Decoding of inconsistent communications. *Journal of personality and social psychology 6*, 1 (1967), 109–114.

11. Peat, J., Elliott, E., Baur, L., and Keena, V. *Scientific Writing, Easy When You Know How*. 2002.

12. Spence, J., Frohlich, D.M., and Andrews, S. Performative experience design. *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13*, ACM Press (2013), 2049.

13. Guide to a Successful Paper or Note Submission « CHI 2016. https://chi2016.acm.org/wp/guide-to-a-successful-paper-or-note-submission/.