A Little Respect: Four Case Studies of HCI's Disregard for Other Disciplines

Joe Marshall
Mixed Reality Lab
School of Computer Science
University of Nottingham
Nottingham, UK
joe.marshall@nottingham.ac.uk

Jocelyn Spence
Mixed Reality Lab
School of Computer Science
University of Nottingham
Nottingham, UK
jocelyn.spence@nottingham.ac.uk

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

Stefan Rennick Egglestone
University of Nottingham
Nottingham, UK
stefan.rennick_egglestone@nottingham.ac.uk

CHI'17 Extended Abstracts, May 06 - 11, 2017, Denver, CO, USA

ACM 978-1-4503-4656-6/17/05…$15.00
DOI: http://dx.doi.org/10.1145/3027063.3052752

Abstract
HCI research often demonstrates lack of respect for other disciplines, evidenced by the way work from those disciplines are cited in CHI papers. We present 4 case studies that demonstrate; 1) that HCI researchers sometimes misunderstand and misrepresent work from other disciplines, and 2) how initial misrepresentations can become 'accepted wisdom 'within HCI. This disregard for other disciplines leads to errors such as authors citing work to support 'facts' precisely opposite to those demonstrated by the cited literature. We conclude with recommendations for authors, editors, publishers and readers on how to reduce the risk of such failures.

Author Keywords
HCI; interdisciplinarity.

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

Introduction
We present here four case studies of failure in HCI. In each study, HCI engages with external disciplines via citation in a way that we argue demonstrates lack of respect for findings of that discipline, by misrepresenting, omitting or misunderstanding the

cited work. As an interdisciplinary team, we describe case studies from our expert areas; we suspect misconceptions also arise in other areas, i.e. that these failures suggest a wider problem amongst HCI researchers who not engage sufficiently critically with work from external fields.

The failures occur primarily in background and motivational sections where related work is most discussed. Whilst often not considered the most important sections of publications, errors in 'why work is done' and in aims of work can (and we believe do) lead to research that will achieve results that support erroneous aims. Such misguided understandings of the foundations of research work may not only affect single pieces of research, as others repeat these mistakes. Whilst large scale mis-citation of prior work occurs in other fields (e.g. [26]); we believe HCI's focus on applied problems and use of research from application areas makes it particularly susceptible to such errors.

Failure 1: Exertion games research misrepresents one child health study

What is misinterpreted

A single study by Vandewater, Shim and Caplovitz [27] (referred to here as VSC) which reports correlations between physical activity/inactivity, obesity and video game use. It is heavily cited in exertion gaming literature. VSC is an observational study with multiple results. In particular, it found no correlation between overall sedentary activity and video game use, and while it found a correlation between obesity and videogames, it was a complex curvilinear correlation: "*Children with higher weight status played moderate amounts of electronic games, while children with lower weight status played either very little or a lot of electronic games*"[27]. VSC is not evidence that reducing video game play is the solution to obesity: "*data available to date do not support the notion that turning off the television or unplugging the video game console amounts to a ''magic bullet'' which will reduce the prevalence of childhood obesity*" [27].

How is it misinterpreted in HCI?

Many exertion games papers claim potential to cure obesity, something we previously [17] argued is implausible. In previous work [15] we found that many exertion gaming obesity claims are founded on the argument that: 1) Inactivity causes obesity, 2) Playing videogames causes inactivity. 3) By replacing videogames with exertion games, this effect will be removed, 4) So exertion games can cure obesity.

Surprisingly, many such arguments are supported by reference to VSC to justify the existence of a causal link between videogames, inactivity, and obesity. This is wrong. Firstly, VSC does not even demonstrate purely positive correlations between obesity and video game use, or inactivity and video game use. Secondly, as an observational study at a single point in time it cannot provide evidence for causality. We surveyed all citations to VSC within exertion gaming research (83 citations) and found that 69% of such citations were wrong: 21 papers implied that VSC found a positive linear correlation between videogames and obesity, e.g. *increased television viewing and video game use is associated with overweight in children"* [12], 15 directly described a purely positive correlation between games and obesity, e.g. *"children who had greater average game-time minutes also had higher BMIs than the children with lower average game-time. "* [21], and 22 papers suggested that VSC demonstrates a causal link

between obesity and videogames, e.g. *these advances in technology have led children to adopt a sedentary behavior, causing an increase in obesity*" [5].

Why we think this happens?
There is a general folk perception that increase in obesity is primarily caused by a decrease in activity. This persists despite research evidence that 1) differences in physical activity may not account for differences in weight [7,28]; 2) rather than inactivity causing obesity, inactivity may rather be a symptom caused by people gaining weight then moving less [8,20]; and 3) physical activity interventions have limited effect without associated dietary interventions [14]. Researchers appear not to seek out evidence that is contrary to pre-existing folk perceptions.

There is an observable spread of this error within exertion games also. VSC is first cited in HCI in 2005. By 2007 it is cited as showing "*strong correlation between video game play and obesity*" [24], implying but not stating positive linear correlation. In 2008 the incorrect correlation is described: "*a higher amount of time playing video games was associated with a higher BMI in children*" [18]; by 2009 it is cited to demonstrate that video games cause inactivity and obesity: "*Video games are considered the main reason for physical inactivity*" [10]. We suggest authors may have seen VSC cited, and instead of reading and critiquing the original, cited it in turn as a convenient prop for their argument, creating a scenario where findings become increasingly distorted over time.

Failure 2: HCI simplifies the concept of 'performance' whilst ignoring implications of performance studies literature

What is misinterpreted
The 'misinterpretation' of (theatrical) performance studies research within the HCI community rarely involves such clear-cut misuse of cited material. Instead, misinterpretations tend to be 'sins of omission'. A publication will briefly introduce concepts relating to performance—perhaps not even from performance literature—and present them as though they are uncontested, integral, and representative of the field. HCI readers have no means of knowing the brief mentions of performance they encounter are at best weak indicators of the wealth of potentially useful research that exists in those fields [25].

Possibly more damaging is the tendency to refer to performance and performativity without explicit definitions. This forces readers to rely on their own assumptions about what the terms are referring to, which can be wildly at odds with what is meant in the source material or by other researchers coming from different traditions. 'Performance' can be taken to refer to the fictional storylines presented in a Shakespeare play, the process of adopting a character other than one's own, the productions of the theatre industry—or interactions that deliberately avoid fictions, characters, and theatre spaces altogether. 'Performativity' complicates the situation even further, implying either fakery or the development of the self, depending on the researcher's perspective. Any number of HCI papers speak of 'performance' or 'performativity' and rely on the reader's assumptions about what they might mean, regardless of the misunderstanding that might arise.

How is it misinterpreted in HCI

A recent paper that commits what we would consider a 'sin of omission' is Chen et al's *From interaction to performance with public displays* [3]. 'Performance' takes pride of place in title, abstract, introduction, discussion, and conclusion, and the authors aim to 'use "performance" as an analytical lens' [4 p. 1617]. They do make reference to an important work from performance studies literature about forming audiences [9] in their section about 'an extended notion of audience' [4p. 1628]. However, they dedicate only two sentences to that article before reverting to (seminal and well-chosen) works from the HCI literature. All other substantial references to performance are devoted to the work of Erving Goffman.

We concur with Chen et al that Goffman inspires a great deal of work in HCI involving performance [4 p. 1617], but we are discouraged by this. For all his many contributions, Goffman was a social scientist who used performance as a metaphor for everyday experience, *not* as an object or methodology for research. Citing a metaphor and claiming therefore to have explored the field that the metaphor refers to is disingenuous, and implies there is little of interest to the HCI community in the field of performance studies itself.

We are not privy to Chen et al.'s intentions or deeper research interests behind the paper we use as an example, so cannot say which strands of performance research they might have benefited from, or which strands they have already rejected for good reason. From our perspective, though, we would suggest to interested readers that there are a number of possibilities from within the performance studies literature for extending the ideas raised in [3], such as [13] regarding connectedness and [6] regarding small group interactions. As it stands, Chen et al seem to imply that performance can be used as the primary analytic lens for an entire project when encapsulated in work of one sociologist and brief mention of one actual performance studies researcher.

Why we think this happens

We want to stress that a large percentage of the HCI research that we would fault in this way is produced by researchers whom we know for a fact have a substantial knowledge of, and respect for, performance studies as a discipline. We certainly do not assume ignorance on anyone's part, especially in a field dominated by the need for concise literature reviews and fact-oriented findings. At a fundamental level, although a great deal of effort has been made by some design researchers to integrate the arts into the HCI community [25], arts-based fields do not easily lend themselves to the scientific methods that HCI is rooted in [1]. However, we believe that those of us who attempt to use performance tools in HCI should present current, relevant research from the actual discipline of performance studies. Otherwise, it is all too easy for references to Goffman's sociological frameworks based on metaphors of performance to get shortened to Goffman's theories of performance, and from there to complete discussions of performance supported primarily by reference to Goffman—the situation that HCI finds itself in at times today.

Failure 3: Our 3rd Case Study in a Previous Version of this Paper

What was misinterpreted?

A previously reviewed and accepted version of this paper had a section relating to research on therapy involving computers. In this, we noted that influential work in the field made arguably over-strong statements as to the evidence for certain forms of therapy. We argued that this had led to misinterpretation by less knowledgeable HCI researchers in work that followed. We also noted what we think are some misconceptions relating to therapy in some of that following work. As consistent with the other case studies, we cited only work we directly quoted from.

How was it misinterpreted in HCI?

The day before the conference, we received upset emails from authors of the influential work cited, suggesting we had written a take-down of their work, which was not our intention. The case study was written unclearly, and because only their work was cited could be interpreted as us arguing these knowledgeable therapy experts had made the errors we described. We immediately contacted the publishers to discuss retracting or revising the publication to avoid creating our own ongoing misconception.

Why do we think this happened?

This happened arguably due to a lack of respect for therapy research – at a point during the writing and revision of this paper, the author expert on therapy changed their job and became less research active. Because of this, some last minute revisions of that section for the submitted version of the paper were done by the first author who has only passing knowledge of the subject, and the nuance was lost.

Failure 4: Overstating the Importance of Non-verbal Communication (The 'Myth of Mehrabian')

What is misinterpreted?

Research by Mehrabian studied inconsistency between spoken words and tone of voice [19]. Participants were shown films of people saying single words chosen to be positive (e.g. "Thanks"), negative ("Brute") or neutral ("Maybe"). They were spoken in negative, positive or neutral tones of voice independent of word. The research showed that for words where tone and content directly conflict, tone of voice was significantly more influential than text content in people's assessment of emotional intention of the speaker.

How is it misinterpreted in HCI?

A prominent folk myth misinterprets this research by Mehrabian and states that 93% of all communication is non-verbal, and only 7% of communication is verbal (as opposed to inconsistent single words without significant textual content as in the study). This widely cited misinterpretation is trivially disproved by the existence of textual forms of communication such as books, email and the internet (see [29] for full debunking of the "Myth of Mehrabian"). It continues to be stated in HCI, for example: "*7% of attitude communication depends on the words spoken*" [23], and "*gestures, facial expressions or the way we use our voice, play a more significant role during an interaction than its verbal counterpart*" [4].

Why we think this happens?

This myth is primarily quoted in affective computing, where it is used to justify research on facial expressions, tone of voice, body language and other non-verbal signals. Arguably, this happens for the same

reason the 7% myth is highly attractive to communications and body language coaches; because it makes an argument that what is studied is highly important. It is undoubtedly more attractive to write that "93% of communication is non-verbal, so it is vitally important to create non-verbal communication behaviors in robots" than it is to write something less urgent such as "non-verbal communication is an important side-channel to communication, so we think robots should make use of it".

Discussion: Causes of these problems

In all our case studies, common threads of work which are poorly founded continue to exist, and to pass peer review. There seems to be no effective mechanism in HCI for catching or fixing these errors and omissions. We suggest that there are several potential causal factors underlying these problems:

Concise HCI writing and citation style

The ACM format used in many HCI publication venues, and fixed length conference publications, gives authors strong incentives to be extremely concise in their writing. Further to this, the numerical referencing means that readers usually cannot infer what a citation is pointing to without cross referencing within the paper. As we describe in [16], this may be exacerbated by the tendency in CHI to cite prior work with either no description, or to cite prior work purely as facts, such as: '*sky is blue[citation]*'. Thus the vast majority of citations are likely to remain unchecked by readers or reviewers, so misleading citation may be hard to spot.

Post-hoc literature review construction

While in theory, understanding the literature may be something that is done prior to starting a project, from our experience as reviewers and writers for CHI, we believe that in practice, background and literature review sections of papers are often written after doing work. Literature reviews are constructed to justify what has already been done on a project, to convince readers of the novelty of that work, or convince reviewers of the existence of research "gaps" rather than to openly seek, understand and critique the best evidence available. Brooks describes this mode of working as "*authors can be pictured as intellectual partisans of their own opinions, scouring the literature for justification*" [2]. As such, only a limited and biased subset of relevant work seems to be reviewed during the process. Further, with such limited literature reviewing, and without coherent discussion of work cited, we risk people simplifying and misrepresenting the research we cite.

Citation Citing

Our first case study suggests a situation where one researcher cites something as fact, then other researchers wishing to demonstrate that fact either cite the same thing as supporting the same fact, or cite the HCI article to support that fact. By the time something has been cited by someone following someone else, the risk is that nuances of the research are completely lost; this is exacerbated by the fact that prior work is rarely even described, let alone critically analyzed [16].

Lack of writer expertise

The highly interdisciplinary nature of HCI has often been criticized for encouraging what Penny calls 'shallow interdisciplinarity' [22], where work is based on references from other disciplines, but does not deeply engage in the actual nature of that work. Our exertion game failure is a classic example of this, in

that rather than engage in what is an extremely complex and as yet uncertain area of health literature when dealing with a multifactorial societal problem such as obesity, authors rely on 'common sense' folk knowledge which vastly oversimplifies the issue, and then delve into the literature just enough to get a citation to support this knowledge.

A further risk to this kind of shallow citation occurs when, as we see with the Vandewater et al., paper, citation occurs in an otherwise interesting paper which is itself well cited. Future authors see a citation in an influential and authoritative paper, assume it is correctly cited and use it to support similar arguments.

Lack of reviewer expertise

While HCI conferences and journals aim for an extremely broad range of reviewers, it is inevitable that there will not be reviewers who are expert in all possible application areas, or in a complete range of theoretical approaches and philosophies. Given front sections of papers are so concise and limited, reviewing is likely to be highly focused on considering other sections of the paper containing descriptions of study, discussion etc. Furthermore, reviewer choice is likely to be biased towards reviewers who can meaningfully critique these 'more important' sections. So work that in its motivation mis-cites public health research, or claims to use performance theory whilst failing to engage with modern conceptions of performance, will often not get picked up on.

Politeness and lack of space for criticism

In other work [16], we present statistics relating to CHI2016 which demonstrate that the prevailing style of citation in CHI is to present prior work without any discussion, in a way that discourages critique or analysis. Whilst there are a small number of papers which take task with the quality of HCI research, such as Greenberg & Buxton's critical discussion of usability evaluation [11] and Spence et al.'s critique of performance in HCI [25], mainstream HCI does not integrate criticism. Further, none of the major journals such as ACM TOCHI, International Journal of Human Computer Studies, or indeed CHI conferences have any format for critical response to published articles, beyond questions at a CHI presentation. There are no 'letters to the editor' sections or critical responses, no culture of pre-prints and early criticism. This means that once a piece of HCI work is in publication, it is unlikely to attract any critical discussion, except possibly in essays such as this one. While we clearly believe essays have significant value, these are unlikely to be seen in a search for related work by someone wishing to do core practical research in HCI.

Conclusions: What Next?

We conclude this paper with 3 suggestions for how CHI authors, reviewers and organisers might work to reduce the risk of failures such as we describe above.

As readers, we should be critical about claims relating to external fields

As an interdisciplinary field that engages with technology use in many areas of society, HCI often requires us to engage with external literature. We must engage critically with that literature itself, rather than simply skim-read sufficiently to support our wider beliefs, or engaging with it in a second hand way via HCI literature. So rather than assuming that a complex multi-factorial societal problem such as obesity is a simple matter of doing a bit more exercise, as we have

read in another HCI paper (see Failure 1 above), we should consider what underlying expert literature actually says about the success of exercise interventions in obesity treatment and consider whether it is the right thing to do. Further, with the growing agenda for HCI to 'do good', such as to engage positively with society, encourage health, and support those in less developed countries, HCI is increasingly engaging with complex situations where, as our exertion game example demonstrates, simplistic folk knowledge explanations of societal issues can lead directly to poor design.

As reviewers and readers we should consider motivation and related work sections

We do not believe major statistical errors, or obvious errors in study methods would propagate through HCI in the same way as our case studies. As reviewers and readers of work, we should consider related work sections beyond the typical level of whether there are missing citations, in particular, reviewers should be encouraged to read through to sources when reviewing when they are unsure of them, something which clearly has not happened in some of our examples.

As organizers/editors we should consider all application domains involved when choosing reviewers

We note that CHI has increasingly created subcommittees for popular application areas such as health and computer gaming. However, as ACs and reviewers for several ACM HCI conferences and journals, we believe that it is still common for papers on the subject of "games for health" work for example to be reviewed only by games reviewers with no expert input from health academics, or for persuasive computing work to be reviewed with no psychologists on the review panel. This inevitably raises the risk of mistakes going un-noticed. Further to this, as reviewers, we should be clear as to which aspects we can and cannot evaluate of work, so for example a games reviewer should be clear that they may not be able to evaluate health elements of games for health.

References

1. Bardzell, J. and Bardzell, S. Cover story: Humanistic HCI. *interactions 23*, 2 (2016), 20–29.
2. Brooks, T.A. Private Acts and Public Objects: An Investigation of Citer Motivations. *Journal of the American Society for Information Science 36*, 4 (1985), 223–29.
3. Chen, J., Dourish, P., Hayes, G.R., and Mazmanian, M. From interaction to performance with public displays. *Personal and Ubiquitous Computing 18*, 7 (2014), 1617–1629.
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. Delden, R. Van, Moreno, A., Ramos, C., Reidsma, D., and Poppe, R. Hang in There : A Novel Body-Centric Interactive Playground. *eNTERFACE 2013, IFIP AICT 425 Y. Rybarcz*, (2014), 160–178.
6. Dolan, J. *Utopia in performance : finding hope at the theater*. University of Michigan Press, 2005.
7. Dugas, L.R., Harders, R., Merrill, S., et al. Energy expenditure in adults living in developing compared with industrialized countries: a meta-analysis of doubly labeled water studies. *American Journal of Clinical Nutrition 93*, 2 (2011), 427–441.
8. Ekelund, U., Brage, S., Besson, H., Sharp, S., and Wareham, N.J. Time spent being sedentary and weight gain in healthy adults: reverse or

bidirectional causality? *The American journal of clinical nutrition 88*, 3 (2008), 612–7.

9. Ginters, L. On Audiencing: The Work of the Spectator in Live Performance. *About Performance*, 10 (2010), 7.
10. Gorgu, L., O'Hare, G.M.P., and O'Grady, M.J. Towards mobile collaborative exergaming. *2nd International Conference on Advances in Human-Oriented and Personalized Mechanisms, Technologies, and Services - CENTRIC 2009*, (2009), 61–64.
11. Greenberg, S. and Buxton, B. Usability evaluation considered harmful (some of the time). *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*, (2008), 111.
12. Henderson, K.E. and Schwartz, M.B. Treatment of overweight children: Practical strategies for parents. In J.D. Latner, G.T. Wilson, J.D. (Ed) Latner and G.T. (Ed) Wilson, eds., *Self-help approaches for obesity and eating disorders: Research and practice.* Guilford Press, 2007, 289–309.
13. Hill, L. and Paris, H. *Performing Proximity: Curious Intimacies*. Palgrave Macmillan, 2014.
14. Johns, D.J., Hartmann-Boyce, J., Jebb, S.A., and Aveyard, P. Diet or Exercise Interventions vs Combined Behavioral Weight Management Programs: A Systematic Review and Meta-Analysis of Direct Comparisons. *Journal of the Academy of Nutrition and Dietetics 114*, 10 (2014), 1557–1568.
15. 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).
16. Marshall, J., Linehan, C., Spence, J., and Rennick Egglestone, S. Throwaway Citation of Prior Work Creates Risk of Bad HCI Research. *In submission, alt.chi 2017*, (2017).
17. Marshall, J., Mueller, F. "Floyd," Benford, S., and Pijnappel, S. Expanding exertion gaming. *International Journal of Human-Computer Studies 90*, (2016), 1–13.
18. Mealey., L.M. Effect of Dance Dance Revolution on energy expenditure and enjoyment in severely overweight children compared to alternative forms of activity. 2008.
19. Mehrabian, a and Wiener, M. Decoding of inconsistent communications. *Journal of personality and social psychology 6*, 1 (1967), 109–114.
20. Metcalf, B.S., Hosking, J., Jeffery, A.N., Voss, L.D., Henley, W., and Wilkin, T.J. Fatness leads to inactivity, but inactivity does not lead to fatness: a longitudinal study in children (EarlyBird 45). *Archives of disease in childhood 96*, 10 (2011), 942–7.
21. Neese, H.M. Factors Influencing the Implementation of Standards-Based Elementary School Physical Education. *Thesis*, (2012).
22. Penny, S. Rigorous Interdisciplinary Pedagogy: Five Years of ACE. *Convergence: The International Journal of Research into New Media Technologies 15*, 1 (2009), 31–54.
23. Robb, D.A., Padilla, S., Methven, T.S., Kalkreuter, B., and Chantler, M.J. A Picture Paints a Thousand Words but Can it Paint Just One? *Proceedings of the 2016 ACM Conference on Designing Interactive Systems - DIS '16*, ACM Press (2016), 959–970.
24. Sinclair, J., Hingston, P., and Masek, M. Considerations for the design of exergames. *Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia - GRAPHITE '07*, ACM Press (2007), 289.

25. 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.
26. Todd, P.A., Guest, J.R., Lu, J., and Chou, L.M. One in four citations in marine biology papers is inappropriate. *Marine Ecology Progress Series 408*, (2010), 299–303.
27. Vandewater, E.A., Shim, M.S., and Caplovitz, A.G. Linking obesity and activity level with children's television and video game use. *Journal of Adolescence 27*, 1 (2004), 71–85.
28. Wilks, D.C., Sharp, S.J., Ekelund, U., et al. Objectively measured physical activity and fat mass in children: a bias-adjusted meta-analysis of prospective studies. *PloS one 6*, 2 (2011), e17205.
29. Yaffe, P. The 7% rule. *Ubiquity 2011*, October (2011), 1–5.