

Creating a User Manual for Healthy Crowd Engagement: A Review of Mark Hedges and Stuart Dunn's *Academic Crowdsourcing in the Humanities: Crowds, Communities and Co-production*

Samantha Blickhan <samantha_at_zooniverse_dot_org>, Zooniverse & Adler Planetarium

Abstract

This piece examines *Academic Crowdsourcing in the Humanities: Crowds, Communities and Co-production*, by Mark Hedges and Stuart Dunn, and considers the role it plays within existing literature about crowdsourcing and digital humanities.

In an academic context, “crowdsourcing” refers to the act of presenting data to members of the public, and inviting them to perform a task, or series of tasks, upon those data. It is a broad definition, but crowdsourcing is a broad methodological concept that can be applied in a number of ways, from academia to business practice. As their title suggests, Mark Hedges and Stuart Dunn narrow the scope considerably with *Academic Crowdsourcing in the Humanities: Crowds, Communities and Co-production*, aiming to offer practical guidance, theoretical frameworks, and real-life examples from crowdsourced humanities research projects. Even within this limited scope, there is a wide range of possibility for crowdsourcing when used as a method of academic research, and the authors structure their text accordingly.

1

Such breadth of methodological usage is not new territory for authors writing about digital projects in the humanities (and it should be said that the focus within this text is overwhelmingly on digital methods, though the authors acknowledge crowdsourcing’s pre-digital roots), which can be difficult to fit within a single framework, particularly due to frequent advances in the technologies that support digital projects. The potential audience for this methodological text can include everyone from participants in digital humanities projects to practitioners — including (but not limited to) researchers, archivists, academics, librarians, institutional staff, and students.

2

The authors address this breadth of scope in content as well as audience by using existing literature from other participatory research fields (most notably the adjacent field of “citizen science”) as a starting point on which to build their framework. They note that, under the umbrella of “citizen science”, projects can either focus on data processing (“delegative” tasks) or on bringing in external participants to the process of research (“democratizing” tasks) [Hedges and Dunn 2017, 4]. Using the framework of citizen science allows the authors to present the parallel outcomes from humanities-focused projects. These include a distinction between “academic knowledge” and the knowledge production typically associated with academic crowdsourcing, as well as frequent outcomes for humanities crowdsourcing, such as content transformation and information synthesis. However, this discussion also allows the authors to identify where the framework of citizen science does not meet the needs of the humanities community, thereby requiring a separate system within which this work can take place [Hedges and Dunn 2017, 8].

3

Creating a framework for any digital research method presents a unique challenge. The speed with which tools are released, adapted, and discontinued — in conjunction with the relatively slow development and practice of the academic research process, particularly in regard to traditional methods of publication — means that any sort of so-called “standard practice” can be very difficult to sustain, and research making use of digital technologies often becomes commonplace within academic fields before any type of structure can be proposed, peer-reviewed, and refined.

4

A lack of standardization in the early stages of implementing new academic research practice does not necessarily guarantee a negative outcome for the work conducted within these nascent technological environments. Such instances of trial and error can allow new practices to emerge, free from traditional boundaries, but it often means that meta-textual studies must simultaneously function as both historiography and theory for these fields: critiquing, reviewing and suggesting best practices for their continued use. Hedges and Dunn have set out to provide just such a structure with *Academic Crowdsourcing in the Humanities*, a work that began as a “typology of arts and humanities crowdsourcing methods” [Hedges and Dunn 2017, xii], but which ultimately functions as a sort of how-to guide for those who wish to create, evaluate, or otherwise engage with academic crowdsourcing in the humanities. Throughout the book, the authors incorporate interviews with experts, practitioners and participants in crowdsourcing projects, gathered from Hedges’ and Dunn’s 2012 *Crowdsourcing Scoping Study* and a PARTHENOS project-supported *Foresight Study*, which bring an appropriately co-created element to the text.

5

Though researchers in the humanities have acknowledged crowdsourcing as a method since the early 2010s, the majority of publications on the subject have been overviews of how crowdsourcing has been used within specific humanistic disciplines or sub-disciplines (examples include [Ferriter 2016] [Ridge 2014]), explorations of the effect of public engagement methods in the digital humanities [Terras 2016], or research studies resulting from individual crowdsourcing projects (see for example any of the articles produced by the *Transcribe Bentham* team, including [Causer and Wallace 2012] and [Causer and Terras 2014]). These early publications are thoroughly referenced in *Academic Crowdsourcing in the Humanities*, and their existence essential to the creation of this text, but Hedges’ and Dunn’s work is filling a much-needed gap in the literature by providing a general framework that will be useful for students, research professionals, and members of the public alike.

6

The first three chapters of the book are devoted to an overview of crowdsourcing as a method, including an examination of the historical and technological developments that set the stage for a rise in modern academic crowdsourcing (including the World Wide Web, Web 2.0, and business-oriented crowd models); the scientific origins of crowdsourcing as a method, and how they have influenced subsequent humanities-based work; and — perhaps most importantly — the presentation of a typology of crowdsourcing, modeled after Short and McCarty’s *methodological commons* for digital research methods in the early 2000s [Hedges and Dunn 2017, 2]. A number of projects are used as examples to illustrate many of the practices discussed in the early chapters, and actual case studies do not appear until Chapter 4, where they are used mainly to further illustrate examples of the typology presented in Chapter 3. This structure allows Hedges and Dunn to present a useful overview of how academic crowdsourcing projects work, without veering into a prescriptive approach.

7

Once the authors have presented the history and methods, and offered a shared vocabulary for the actions being taken within digital crowdsourcing projects, they turn their attention to the social elements of crowdsourcing, both for individuals and communities alike. Topics include roles within projects (Chapter 5), motivations for and benefits of participation (Chapter 6), ethical issues like exploitation of volunteer labor and questions of data ownership (Chapter 7), and the role crowdsourcing can play in the creation of knowledge and memory in regard to cultural heritage (Chapter 7). The chapter on ethics is particularly welcome. When digital methodologies involve members of the public in academic work, as is the case with crowdsourcing, it becomes essential to create a system within which projects can be evaluated, as such systems are needed not only for the benefit of the academic research output, but to ensure the fair and ethical treatment of the communities participating. There has been a fair amount of work published on the ethics of paid crowdsourcing, in academia and other platforms like Amazon’s Mechanical Turk (see for example [Williamson 2016]), but this is not the case regarding publications on unpaid academic crowdsourcing.

8

The discussion of labor and exploitation within the chapter on ethical issues presents commonly-asked questions about the ethical grey area that exists around volunteer participation in academic research [Hedges and Dunn 2017, 108–113]. In this section, the interview excerpts are useful for illustrating some of the questions which do not have a definitive answer, especially those which relate to the various justifications presented for why academic crowdsourcing should not be considered unethical. Hedges and Dunn do not shy away from these often difficult topics, and the way they present this information gives a unique insight into the care with which project creators must go about the work of designing, developing and running projects: “While participants do not in general regard themselves as exploited, their

9

willingness to volunteer and their professed enjoyment in participating does not in itself imply that humanities crowdsourcing is in ethical terms positive, or even neutral” [Hedges and Dunn 2017, 113].

For all cases involving ethics and crowdsourcing, the authors note that transparency and regular communication are requisite elements for projects involving the public, as well as acknowledgment of participants and open access to project outcomes in the form of data. The suggestion of access to outcomes in the form of data is critical, but the authors miss the opportunity to discuss a delicate but important issue: that of open access to resulting publications which make use of project data. The authors come close to this topic in an earlier discussion of motivations and benefits, stating “Most humanities crowdsourcing projects, however, do not reward their contributors in material or professional ways”, having listed publication as a material outcome for researchers in the previous sentence [Hedges and Dunn 2017, 91], but fall short of calling for open access publication as a requirement for crowdsourcing project practitioners. In many ways this is part of a larger conversation about access to publication within academia and is certainly not exclusive to the field of crowdsourcing, but it is particularly relevant to a field which invites — and relies upon — public participation. Is it ethical to ask participants to volunteer their time for research, and then put a paywall between those same people and their access to the resulting publications?

10

In the final chapter, “Crowds past, present and future”, the authors reiterate the main arguments from the first eight chapters, and present some suggestions that will benefit practitioners of and participants in digital crowdsourcing in the future. These include human-computer optimization methods, data literacy initiatives for the public, and adoption of open data frameworks (another place where a slight push into adjacent issues of open access publication would have been welcome). Hedges and Dunn note that many projects and wider initiatives already exist which make use of and promote these approaches, but they would certainly benefit from wider adoption.

11

In their conclusion, the authors note that *Academic Crowdsourcing in the Humanities* will not be a step-by-step guide for crowdsourcing practitioners, but is instead meant to provide a theoretical framework to ensure that there is a healthy balance between the quality of project outcomes, and the experience of participants. It is indeed a welcome addition to the growing corpus of publications related to academic crowdsourcing, and when used in conjunction with the existing literature will be a wonderful tool for participants and practitioners alike.

12

Works Cited

- Causer and Terras 2014** Causer, T. and Terras, M. “Crowdsourcing Bentham: Beyond the Traditional Boundaries of Academic History”, *International Journal of Humanities and Arts Computing* 8.1 (2014): 46-64. DOI: <https://doi.org/10.3366/ijhac.2014.0119>.
- Causer and Wallace 2012** Causer, T. and Wallace, V. “Building a volunteer community: results and findings from *Transcribe Bentham*”, *Digital Humanities Quarterly*, 6.2 (2012). Available at: <http://www.digitalhumanities.org/dhq/vol/6/2/000125/000125.html>.
- Ferriter 2016** Ferriter, M. “Inviting Engagement, Supporting Success: How to Manage a Transcription Center”, *Collections: A Journal for Museum and Archives Professionals*, 12.2 (2016): 97-116. DOI: <https://doi.org/10.1177%2F155019061601200204>.
- Hedges and Dunn 2012** Hedges, M. & Dunn, S. *Crowd-Sourcing Scoping Study: Engaging the Crowd with Humanities Research*. Arts and Humanities Research Council (2012).
- Hedges and Dunn 2017** Hedges, M. and Dunn, S., *Academic Crowdsourcing in the Humanities: Crowds, Communities and Co-production*. Elsevier, Inc. (2017).
- Ridge 2014** Ridge, M. (ed), *Crowdsourcing Our Cultural Heritage*. Ashgate, Surrey (2014).
- Terras 2016** Terras, M. “Crowdsourcing in the Digital Humanities”. In S. Schreibman, R. Siemens, and J. Unsworth (eds), *A New Companion to Digital Humanities*, Wiley-Blackwell (2016), pp. 420-439. DOI: <https://doi.org/10.1002/9781118680605.ch29>.
- Williamson 2016** Williamson, V. “Can crowdsourcing be ethical?”, *TechTank* blog (Feb. 3, 2016), <https://www.brookings.edu/blog/techtank/2016/02/03/can-crowdsourcing-be-ethical-2/>.