

Building A Volunteer Community: Results and Findings from *Transcribe Bentham*

Tim Causer <t_dot_causer_at_ucl_dot_ac_dot_uk>, Bentham Project, University College London

Valerie Wallace <valwall23_at_gmail_dot_com>, Bentham Project, University College London, and Center for History and Economics, Harvard University

Abstract

This paper contributes to the literature examining the burgeoning field of academic crowdsourcing, by analysing the results of the crowdsourced manuscript transcription project, *Transcribe Bentham*. First, it describes how the project team sought to recruit volunteer transcribers to take part, and discusses which strategies were successes (and which were not). We then examine *Transcribe Bentham*'s results during its six-month testing period (8 September 2010 to 8 March 2011), which include a detailed quantitative and qualitative analysis of website statistics, work completed by the amateur transcribers, as well as the demographics of the volunteer base and their motivations for taking part. The paper concludes by discussing the success of our community building with reference to this analysis. We find that *Transcribe Bentham*'s volunteer transcribers have produced a remarkable amount of work – and continue to do so, carrying out the equivalent labour of a full-time transcriber – despite the nature and complexity of the task at hand.

Introduction

Crowdsourcing is an increasingly popular and attractive option for archivists, librarians, scientists, and scholarly editors working with large collections in need of tagging, annotating, editing, or transcribing. These tasks, it has been argued, can be accomplished more quickly and more cheaply by outsourcing them to enthusiastic members of the public who volunteer their time and effort for free [Holley 2010].^[1] Crowdsourcing also benefits the public by making available and engaging volunteers with material hitherto only accessible to diligent researchers, or with sources previously considered too complex for a non-expert to understand. A project like *Galaxy Zoo*, for example, has successfully built up a community of more than 200,000 users who have classified over 100 million galaxies, thus supporting a great deal of academic research [Raddick et al 2010]. Crowdsourcing aims to raise the profile of academic research, by allowing volunteers to play a part in its generation and dissemination.

The Bentham Project at University College London (UCL) sought to harness the power of crowdsourcing to facilitate the transcription of the manuscript papers of Jeremy Bentham (1748-1832), the great philosopher and reformer. The purpose of the Bentham Project is to produce the new authoritative scholarly edition of *The Collected Works of Jeremy Bentham*, which is based in large part on transcripts of the vast collection – around 60,000 folios – of Bentham manuscripts held by UCL Special Collections.^[2] The Bentham Project was founded in 1958, and since then 20,000 folios have been transcribed and twenty-nine volumes have been published. The Project estimates that the edition will run to around seventy volumes; before the commencement of *Transcribe Bentham* around 40,000 folios remained untranscribed.

This new edition of Bentham's *Collected Works* will replace the poorly-edited, inadequate and incomplete eleven-volume edition published between 1838 and 1843 by Bentham's literary executor, John Bowring [Schofield 2009a, 14–15, 20–22]. The Bowring edition omitted a number of works published in Bentham's lifetime, as well as many substantial works which had not been published but which have survived in manuscript; a forthcoming *Collected Works* volume

1

2

3

entitled *Not Paul, but Jesus* – only a part of which was previously published by Bentham, and was left out of the Bowring edition altogether – will recover Bentham's thinking on religion and sexual morality. This material has significant implications for our understanding of utilitarian thought, the history of sexual morality, atheism, and agnosticism. Bentham's writings on his panopticon prison scheme still require transcription, as do large swathes of important material on civil, penal, and constitutional law, on economics, and on legal and political philosophy. In short, while Bentham's manuscripts comprise material of potentially great significance for a wide range of disciplines, much of the collection – far from being even adequately studied – is virtually unknown. A great deal of work, both in exploring the manuscripts and producing the *Collected Works*, clearly remains to be done.

The Bentham Papers Transcription Initiative – *Transcribe Bentham* – was established to quicken the pace of transcription, speed up publication of the *Collected Works*, create a freely-available and searchable digital Bentham Papers repository, and engage the community with Bentham's ideas at a time when they are of increasing contemporary relevance.^[3] *Transcribe Bentham* crowdsources manuscript transcription, a task usually performed by skilled researchers, via the web to members of the public who require no specialist training or background knowledge in order to participate. The project team developed the "Transcription Desk", a website, tool and interface to facilitate web-based transcription and encoding of common features of the manuscripts in Text Encoding Initiative-compliant XML. Transcripts submitted by volunteers are subsequently uploaded to UCL's digital repository, linked to the relevant manuscript image and made searchable, while the transcripts will also eventually form the basis of printed editions of Bentham's works.^[4] The products of this crowdsourcing will thus be utilised for both scholarly and general access purposes. *Transcribe Bentham* was established and funded under a twelve-month Arts and Humanities Research Council Digital Equipment and Database Enhancement for Impact (DEDEFI) grant. The funding period was divided into six months of development work from April 2010, and the Transcription Desk went live for a six-month testing period in September of that year.^[5]

Crowdsourcing is becoming more widespread, and thus, it is important to understand exactly how, and if, it works. It is a viable and cost-effective strategy only if the task is well facilitated, and the institution or project leaders are able to build up a cohort of willing volunteers. Participant motivation in crowdsourcing projects is, therefore, attracting more focused attention. The teams behind the *Zooniverse* projects have analysed the motivations and demographic characteristics of their volunteers in an attempt to understand what drives people to participate in online citizen-science projects, while the *North American Bird Phenology Programme*, established to track climate change by crowdsourcing the transcription of birdwatchers' cards, also assessed its participants' opinions. Rose Holley has also offered several invaluable general insights on user motivation, drawing on the experience of crowdsourcing the correction of OCR software-generated text of historic newspapers at the National Library of Australia (NLA), while Peter Organisciak has provided a useful analysis of user motivations in crowdsourcing projects. [Raddick et al 2010] [Romeo and Blaser 2011] [Phenology Survey 2010] [Holley 2009] [Holley 2010] [Organisciak 2010].

However, *Transcribe Bentham* differs from previous crowdsourcing and community collection schemes, in that its source material is a huge collection of complex manuscripts. Though several projects have crowdsourced manuscript transcription, the material they have made available is generally formulaic, or at least reasonably straightforward to decipher and understand [Old Weather Project] [Family Search Indexing] [War Department Papers]. Transcribing the difficult handwriting, idiosyncratic style, and dense and challenging ideas of an eighteenth and nineteenth-century philosopher is more complex, esoteric, and of less immediate appeal than contributing to a genealogical or community collection.

This paper describes how the *Transcribe Bentham* team sought to attract volunteer transcribers and build an online community. It outlines which strategies worked and which did not, and, drawing on qualitative and quantitative data, analyses the complexion of our volunteer base, comparing its demographic and other characteristics with those of other crowdsourcing projects. This evidence will shed more light on the nature of user participation, crowdsourced manuscript transcription, and provide guidance for future initiatives. Section one will describe our attempts to recruit a crowd and build a community of users; section two will analyse the make-up of this user base, and assess site statistics, user contributions, and motivations; and section three will consider the success of our community building with reference to

4

5

6

7

this analysis.

Crowd or Community?

Caroline Haythornthwaite has discerned two overlapping patterns of engagement in online “peer production” initiatives like *Transcribe Bentham*, distinguishing between a “crowd” and a “community”. Contributions made by a crowd, which Haythornthwaite describes as “lightweight peer production”, tend to be anonymous, sporadic, and straightforward, whereas the engagement of a community, or “heavyweight peer production”, is far more involved. A community of volunteers engaged in the latter requires, Haythornthwaite suggests, qualitative recognition, feedback, and a peer support system. Contributors tend to be smaller in number, to be less anonymous, and to respond to more complex tasks and detailed guidelines. Heavyweight peer production might also involve a multi-tiered progress system to sustain motivation; a crowd on the other hand, is satisfied with quantitative recognition, perhaps in the form of progress statistics, and a two-tiered hierarchy such as that of contributor and moderator. These two patterns, Haythornthwaite contends, are often discernable within one project [Haythornthwaite 2009].

Transcribe Bentham blends both heavyweight and lightweight peer production. We attracted an anonymous crowd of one-time or irregular volunteers, along with a smaller cohort of mutually supportive and loyal transcribers. We aimed to cast our net wide by opening the Transcription Desk to all, by creating as user-friendly an interface as possible, and by simplifying the transcription process as much as we could (Figure 1). But, as transcribing Bentham’s handwriting is a complex and time-consuming task which requires considerable concentration and commitment, we also tried to build a dedicated user community to enable sustained participation by, for example, implementing a qualitative and quantitative feedback and reward system. The following section will describe the strategies we devised first to recruit the crowd, and then to foster the community.

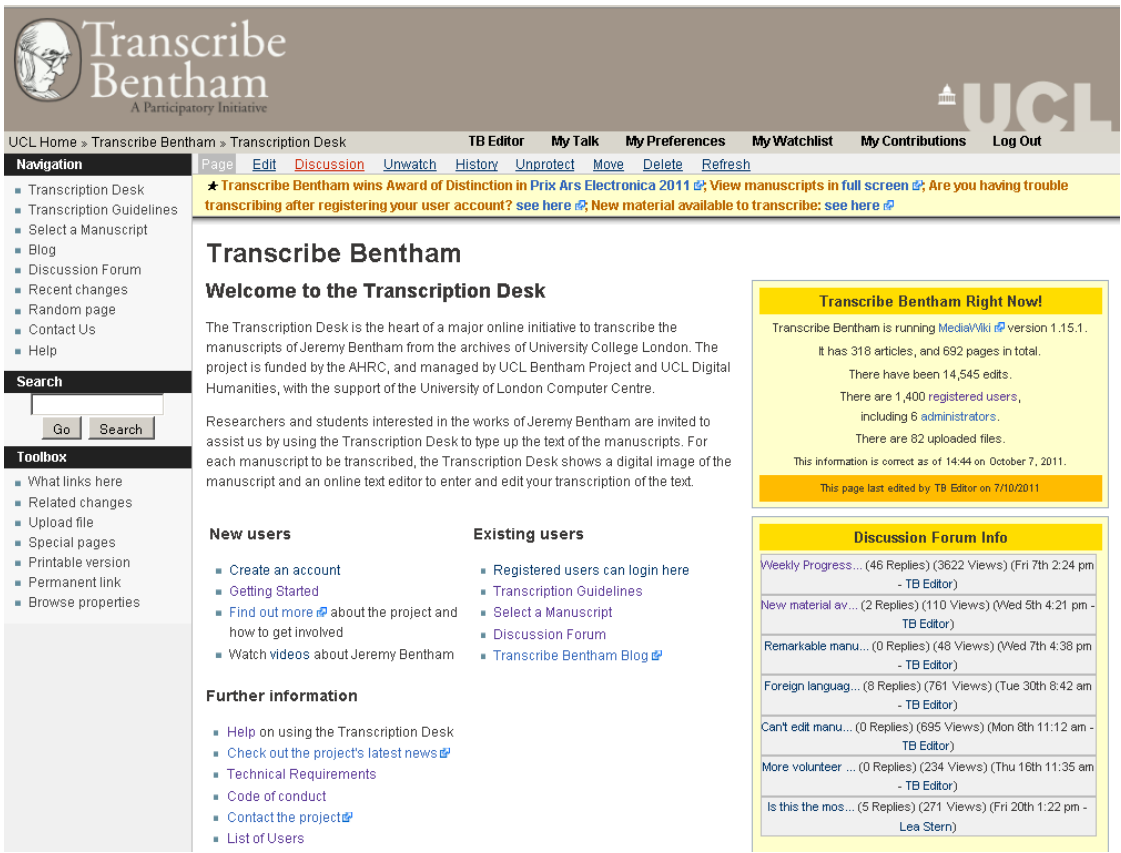


Figure 1. The *Transcribe Bentham* Transcription Desk

Recruiting the Crowd: The Publicity Campaign

Our publicity campaign targeted a variety of audiences including the general public, academic community, libraries, archives professionals, and schools. We devised audience-specific tactics as well as more general strategies, taking advantage of services offered by UCL to help us implement our campaign; these included the various Media Relations, Corporate Communications, Outreach, Public Engagement, and Learning and Media Services teams. In devising these strategies we had to consider issues of cost and timing. *Transcribe Bentham* had a limited budget to spend on publicity – £1,000 – and, as our testing period was six months only, a short time-frame in which to execute the plan. Though we hoped to target the English-speaking world, many of our strategies were, by necessity, confined to the United Kingdom.

10

The General Public

As a web 2.0 project, it was vital to have a visible and interactive online presence. We created a project blog which was regularly updated with progress reports, details of media coverage, and forthcoming presentations, and which linked directly to, and became the main entry point for, users to the Transcription Desk. We also utilised social media by creating a Twitter profile and a Facebook page, which were integrated into our blog and main Bentham Project website, which also prominently featured *Transcribe Bentham*. A Google Adwords account was created when the Desk went live in order to generate traffic which, owing to budget constraints, was established on a trial basis. We prepaid £60 on our account which was exhausted by the end of September.

11

Besides the web we attempted to generate awareness of the project through traditional media. With the help of UCL Media Relations, a press release was drawn up at the launch of the project in September 2010, and distributed to major British newspapers and magazines. UCL Corporate Communications assisted in designing a *Transcribe Bentham* leaflet for distribution, 2,500 copies of which were printed at a cost of £295 (excl. VAT). We distributed the leaflet at academic conferences and institutions in Britain, Europe, and Australasia, and the leaflet was also dispensed throughout the year at Bentham's Auto-Icon in UCL's South Cloisters.^[6] *Transcribe Bentham* was also promoted via a video produced by UCL Media Services, which was embedded into our websites, and hosted on UCL's YouTube channel.^[7]

12

The Academic and Professional Community

At the outset, we believed that the academic and professional community would be the most receptive to our project. We targeted not just existing Bentham scholars, academics, and students with an interest in history and philosophy, but also those interested in digital humanities and crowdsourcing, palaeography, and information studies. We hoped to encourage a range of scholars to embed *Transcribe Bentham* in teaching and learning, thereby helping to build a dedicated user base and encourage Bentham scholarship.

13

We considered placing advertisements in academic journals and more mainstream subject-focused magazines. However, an advert in a single journal, with a limited print run, would have swallowed up nearly half of the publicity budget, and we felt that free coverage in the national press would achieve greater impact. Therefore, in order to reach a potentially diverse academic audience, the press release was sent to, amongst others, *The Guardian*, *TechCrunch*, *The Register*, *Wired*, *Mashable*, *Times Higher Education*, *Times Educational Supplement*, *The Times*, *BBC History Magazine*, and *History Today*. In July 2010 two articles mentioning *Transcribe Bentham* had appeared in the *Times Higher Education*, and it thus seemed sensible to approach that publication in the hope that it would run a follow-up piece [Mroz 2010] [Cunnane 2010].

14

Notifications were sent to a large number of academic and professional mailing lists, online forums, and the websites of academic societies. Though some bodies failed to respond, most of those contacted circulated an announcement about *Transcribe Bentham* via their list or featured it on their websites. Besides these initiatives, project staff delivered presentations on *Transcribe Bentham* at several seminars, conferences, and workshops throughout the year.^[8] We also engaged in consultation with representatives from different repositories including the National Library of the Netherlands, The National Archives (UK), the Natural History Museum (UK), and Library and Archives Canada.

15

To promote *Transcribe Bentham* to palaeography, information studies, and research methods students, we contacted individual academics, libraries, archives, and educational bodies including the Higher Education Academy History Subject Centre, Senate House Library, and The National Archives. This outreach was generally successful and met with

16

enthusiastic responses, though The National Archives responded negatively, stating that only notifications relating to “government departments, archives and organisations directly relevant to the activities of The National Archives” could be posted on their site. On the recommendation of the HEA History Subject Centre, we created pages on the reading of historical manuscripts to demonstrate how *Transcribe Bentham* could be used as a tool in teaching and learning.^[9] The Subject Centre subsequently produced a review of the resource recommending its use for palaeographic and historical training in undergraduate History classes [Beals 2010]. Dr Justin Tonra, then a Research Associate on the project, also contributed a tutorial using *Transcribe Bentham* to *TEI by Example*, an online resource run by the Royal Academy of Dutch Language and Literature, King's College London, and UCL.^[10]

Schools

At the development stage, project members anticipated that school pupils and their teachers, particularly those undertaking A-levels in Religious Studies, Philosophy, History, Law, and Politics, could be another potential audience, especially considering Bentham specifically features in the curricula for Religious Studies and Philosophy. Once the project got underway, it was tailored in such a way so as to attract schools and colleges. We created pages with information explaining how *Transcribe Bentham* related to relevant A-Levels and Scottish Highers, including reading lists and direct links to groups of manuscripts of relevance to particular areas of study.^[11] We aimed, moreover, to target school teachers and pupils through the media and the web. Our press release was sent to educational publications, while notices and invitations to post links to our site were sent to a range of educational websites and bodies.

A-level pupils from the Queen's School in Chester visited the Bentham Project in summer 2010 before the Transcription Desk went live, where they tested the website; their experience was written up in the school's website and in the local newspaper [Chester Chronicle 2010]. A link to the Transcription Desk was later included on the school's virtual learning environment. In order to attract more schools to the project we invited school groups to visit the Project to see the Auto-Icon, hear a short lecture, and participate in the transcription exercise. We drew up a letter outlining these details which we sent, along with the *Transcribe Bentham* leaflet, to c.500 state schools in London, the cost of printing and postage for which was around £360. Raines Foundation School in Bethnal Green, London, responded positively to the outreach letter and arranged a visit in November of A-level Philosophy students who participated in the initiative [Bennett 2010a]. The class teacher and one of his pupils were also interviewed about *Transcribe Bentham* for a broadcast journalism project at City University London.

Success?

In terms of raising awareness of the initiative, the publicity campaign has been a success. Despite mainly targeting English-speakers and the UK, particularly with our press release, the project has received media coverage in twelve countries including the United States, Australia, Japan, Germany, Norway, Sweden, Austria, and Poland. We estimate that the project has been mentioned in around seventy blogs, thirteen press articles, and two radio broadcasts. As of 3 August 2012, we have acquired 853 followers on Twitter, and 339 fans on Facebook. *Transcribe Bentham* has certainly made an impact on the academic community and libraries and archives profession; its progress has been tracked by JISC and the Institute of Historical Research, and it has been reviewed by the Higher Education Development Association, and the Higher Education Academy [Dunning 2011] [Winters 2011] [Elken 2011] [Beals 2010]. *Transcribe Bentham* is also being used as a model for archives discovery by repositories in Europe and North America, and has been featured in the professional blog of the British Library [Shaw 2010]. The project has been embedded, moreover, into teaching and learning in Queen's University Belfast, Bloomsburg University, the University of Virginia, and King's College London.

More recently, *Transcribe Bentham* was honoured with a highly prestigious Award of Distinction in the Digital Communities category of the 2011 Prix Ars Electronica, the world's foremost digital arts competition, and staff were given the opportunity to speak about the project at that year's Ars Electronica festival.^[12] This is testimony to the project's international impact, both inside and outside the academy, with the Digital Communities jury commending *Transcribe Bentham* for its “potential to create the legacy of participatory education and the preservation of heritage or an endangered culture” [Achaleke et al 2011, 206]. *Transcribe Bentham* was also one of five crowdsourcing projects

shortlisted for the 2011 Digital Heritage Award, part of that year's Digital Strategies for Heritage Conference.^[13]

We hoped that our considerable efforts in publicising the project, and crowdsourced transcription, would help us to recruit a large crowd of volunteers. We also implemented strategies to retain this crowd and transform it into a loyal community.

Building the Community

The Interface

Retaining users was just as integral to the project's success as recruiting them in the first place. It was therefore important to design a user-friendly interface which facilitated communication in order to keep users coming back to the site, and to develop a sense of community cohesion [Causar, Tonra and Wallace 2012]. The Transcription Desk was developed using MediaWiki, an interface familiar to, and easily navigable by, the millions of those who have browsed, used and contributed to Wikipedia. It was decided that offering remuneration for contributions would be contrary the collaborative spirit of the project, and so platforms such as Amazon's *Mechanical Turk* were discounted, in favour of open source software. An alternative approach would have necessarily limited participation in *Transcribe Bentham*, as well as the level of engagement with and access to material of national and international significance.

Button																
Function	Line Break	Page Break	Heading	Paragraph	Addition	Deletion	Questionable Reading	Illegible Text	Marginal Note	Underline	Superscript	Unusual Spelling	Foreign Language	Ampersand	Long Dash	User Comment
Rendering	-	-	text	-	text	text	text[?]	[...]	text	text	te ^x t	text	text	&	—	-

Figure 2. The Transcription Toolbar

Editing JB/002/504/001

```
[[File:JB_002_504_001.jpg|500px|right]]
'''[[fullurl:JB/002/504/001?action=edit]] Click Here To Edit]]'''
<!-- ENTER TRANSCRIPTION BELOW THIS LINE -->

<head>18</head> <!-- 17 overwritten -->

<p>[17]<!-- 16 is overwritten -->[chi rend="underline">gap/> to the Fund
&emp;c.</h1> From this mode of
<lb>
appropriation, results a wonderful degree of simplicity <lb>and facility to
the whole business. No new
<lb>
Fund created &#x2014; no new Taxes imposed for the payment <lb>of these
Annuitants: <del>since no source is</del>
<lb>
the <del>great</del>+add> existing fund &#x2014; the Consolidated Fund &#x2014;
<lb>
series without addition: since no sooner is it
<lb>
<del>the</del>+burthened</sic> with one hand, <del>gap/><!-- symbol not clear --> than it
is exonerated &#x2014;
<lb>
and exonerated to a much greater degree than
<lb>
it is <del>the</del>+burthened</sic> &#x2014; with the other. To <del>the</del>+to</del> sell
<lb>
new Annuitants with one hand &#x2014; to buy in the old
<lb>
ones with the other &#x2014; <del>the</del>+such is the process</del> &#x2014; to sell
the new ones dear
<lb>
and buy <del>the</del>+up</del> in the old ones cheap &#x2014; such is the
<lb>
profit &#x2014; of the business. <del>the</del>+superscript>+</h1> <del>the</del>+chi
rend="superscript">+</h1> Reference for other reasons.
<del>the</del>+Studies</unclear> if <del>the</del>+tk</del>+</note>
</p>

<!-- DO NOT EDIT BELOW THIS LINE -->
[[Metadata:({{PAGE}})]]
```

Figure 3. Transcribing and Encoding a Manuscript

The features of MediaWiki were utilised in an attempt to forge a virtual community engaged in heavyweight peer production. We provided detailed, clearly-written guidelines to explain the process of transcription and encoding, along with a “quick-start” guide to summarise the main points. Training videos and downloadable files were embedded in order to provide an audiovisual aspect to the learning experience, and an intuitive toolbar was developed so that volunteers otherwise unfamiliar with text encoding could add the relevant TEI-compliant XML tags at the click of a button (Figures 2 and 3). In order to give regular feedback to users and to provide a platform for shared resources, we included a discussion forum on the Desk's main page where volunteers could swap ideas, ask questions, or make requests of the project editors. Each registered participant was given a social profile which could be left anonymous or populated with an avatar and personal information, including his/her home town, occupation, birthday, favourite movies,

and favourite Bentham quotation (Figure 4). Each volunteer profile also included a personal message board and an “add friends” function; we hoped that registered users would be able to message each other privately or publicly and build up a cohort of transcriber friends.

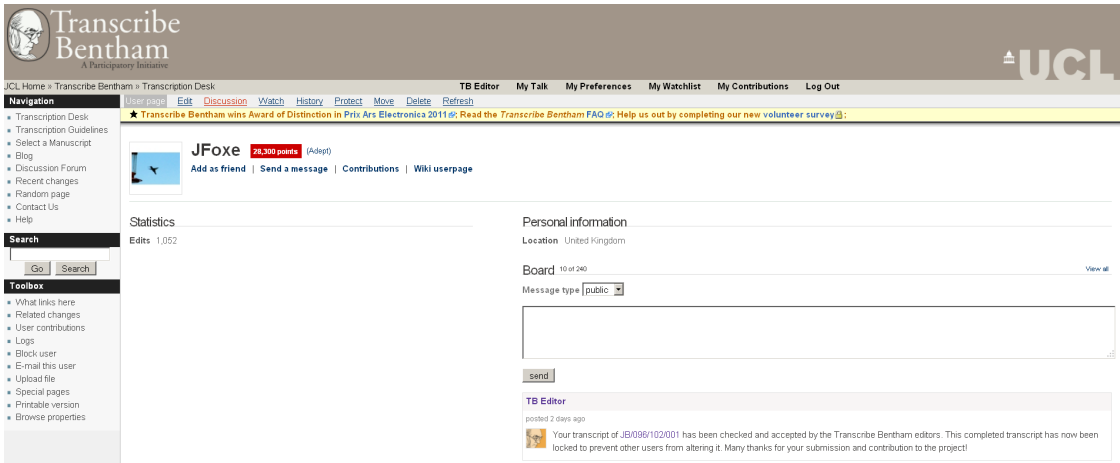


Figure 4. Example of a *Transcribe Bentham* volunteer profile

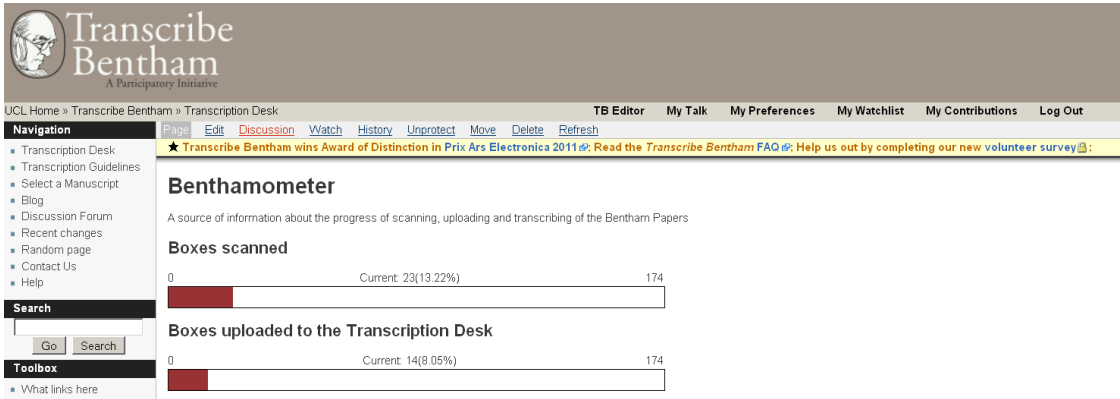


Figure 5. The *Transcribe Bentham* Benthamometer progress bar

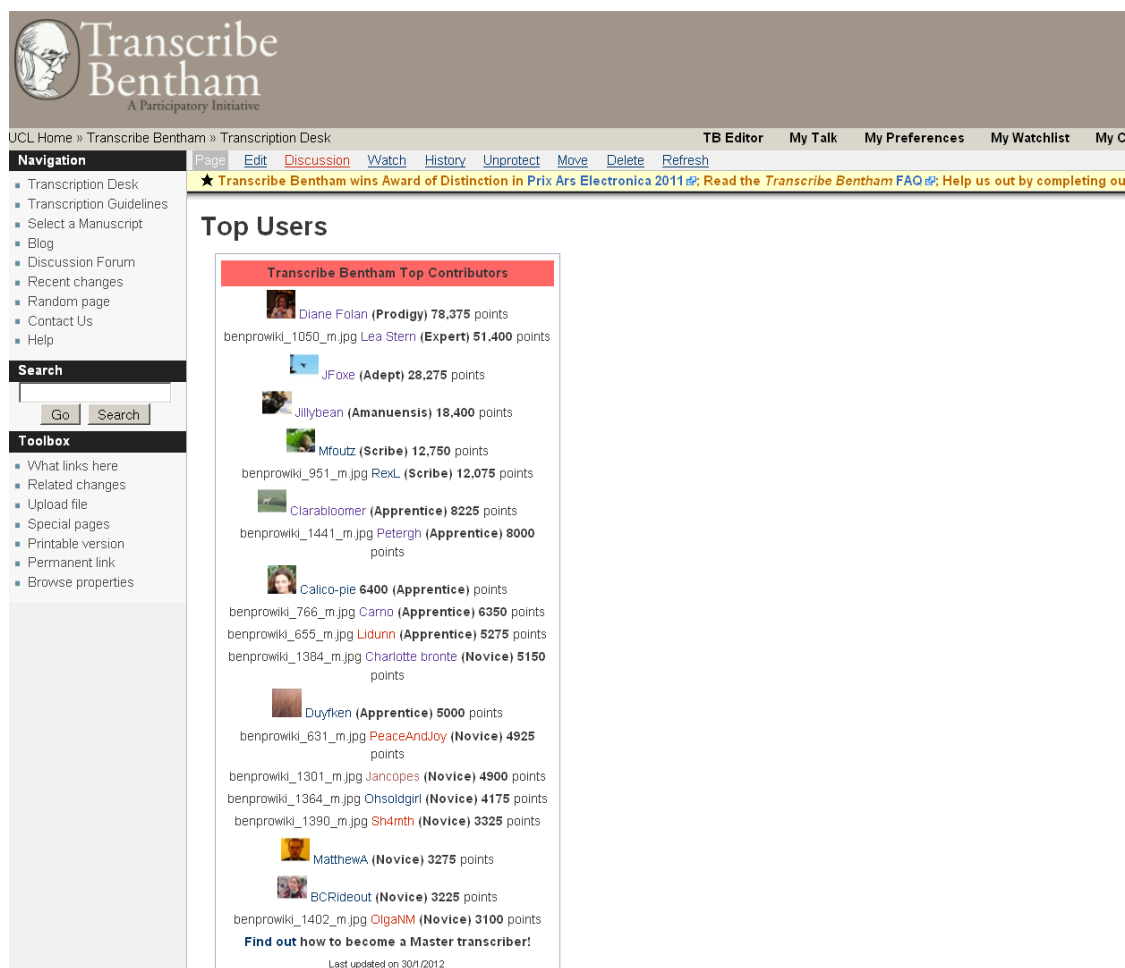


Figure 6. The *Transcribe Bentham* Leaderboard

The project editors used the message function on a daily basis to communicate with and provide feedback to transcribers. The “Benthamometer”^[14] tracked the progress of transcription, while the leaderboard recorded and publicly recognised the efforts of the most diligent transcribers (Figures 5 and 6).^[15] Volunteers received points for every edit made; as an incentive we devised a multi-tiered ranking system, a progress ladder stretching from “probationer” to “prodigy” for transcribers to climb.^[16] We also intended to utilise a gift function which allowed editors to award users with virtual gifts – an image of the *Collected Works* for example – whenever they reached a milestone. “Team-building” features like these have been found to be useful in stimulating participation by other projects like *Solar Stormwatch* and *Old Weather*: we hoped to facilitate interaction between users, to generate healthy competition, and to develop a sense of community. However, some of the social features of the site, including the “add friends” option and the gift-awarding feature, malfunctioned at the development stages. These problems, as will be discussed below, may have been an impediment to social integration.

Community Outreach: Beyond the Virtual

Though we aimed to create a cohesive online community, we were also keen to move beyond the virtual and add a personal element to the initiative by organising a series of public outreach events. This programme was arranged in consultation with local amateur historians and aimed to start a dialogue between professionals and amateurs, to engage the public, and to situate Bentham and UCL more firmly within the local community. We wanted to engage the interest of amateur historians in *Transcribe Bentham* as well as to give our regular transcribers a chance to meet project staff. These events were held in May 2011 and included two information sessions, one held at UCL and one held externally, as well as a guided walk around Bentham’s London.^[17] In terms of integrating the *Transcribe Bentham* community, this strategy, discussed in more detail below, had limited success.

The project team devised, therefore, a range of strategies to recruit a crowd and build a cohort of dedicated transcribers; on his blog discussing crowdsourced manuscript transcription, Ben Brumfield commented that *Transcribe Bentham* “has done more than any other transcription tool to publicize the field” [Brumfield 2011]. As *Transcribe Bentham*'s attempts to crowdsource highly complex manuscripts are novel, the project team was only able to draw on the general experiences of other crowdsourcing projects when making its decisions regarding the recruitment plan. The strategies employed were to a large extent experimental. The following sections of this paper will assess the complexion of our user base and consider how successful these strategies were in forging a *Transcribe Bentham* community.

26

The Results

Our six-month testing period lasted from 8 September 2010 to 8 March 2011, and during this time 1,207 people registered an account (discounting project staff, and seven blocked spam accounts).^[18] Between them these volunteers transcribed 1,009 manuscripts, 569 (56%) of which were deemed to be complete and locked to prevent further editing. Though the fully-supported testing period has ceased, the Transcription Desk will remain available dependent on funding, and *Transcribe Bentham* has become embedded into the Bentham Project's activities. As of 3 August 2012, the project now has 1,726 registered users. 4,014 manuscripts have been transcribed, of which 3,728 (94%) are complete and locked to prevent further editing. However, unless otherwise stated, the analysis below pertains to the six-month testing period.^[19]

27

In this section, we will assess site statistics, user demographics, behaviour, and motivations. Our findings are derived from quantitative data provided by a Google Analytics account,^[20] analysis of statistics collated from the Transcription Desk, qualitative findings from a user survey, and comparisons with other studies of crowdsourcing volunteer behaviour.^[21] The survey received 101 responses – about 8% of all registered users – 78 of which were fully completed. While it is, therefore, not necessarily representative of the entire user base, the survey contains a great deal of revealing information about those who did respond.

28

Before reviewing the data, it is worth taking note of the following milestones in the project's life during the testing period:

29

- 8 September 2010: official launch of the Transcription Desk and first wave of publicity
- 27 December 2010: *New York Times* article featuring *Transcribe Bentham* published online [Cohen 2010]
- 28 December 2010: *New York Times* article published in print
- 1 February 2011: first broadcast of *Deutsche Welle World* radio feature^[22]
- 1 and 2 February 2011: each registered user received an invitation to take part in the *Transcribe Bentham* user survey
- 8 March 2011: end of testing period

As will be seen, the publication of the article in the *New York Times* (NYT) had a vital and enduring impact upon *Transcribe Bentham*, and it is thus helpful to consider the testing period as having had two distinct parts: Period One, or the pre-NYT period, covers 8 September to 26 December 2010 (110 days); and Period Two, the post-NYT period, encompasses 27 December 2010 to 8 March 2011 (72 days).

30

Site Visits

During the six months as a whole, the Transcription Desk received 15,354 visits from 7,441 unique visitors, or an average of 84 visits per day (see Figure 7).^[23] Period One saw 5,199 visits, while in Period Two, there were 10,155. It is clear, then, that traffic to the site during the shorter Period Two was much greater than the longer Period One, but this is far from the full story.

31

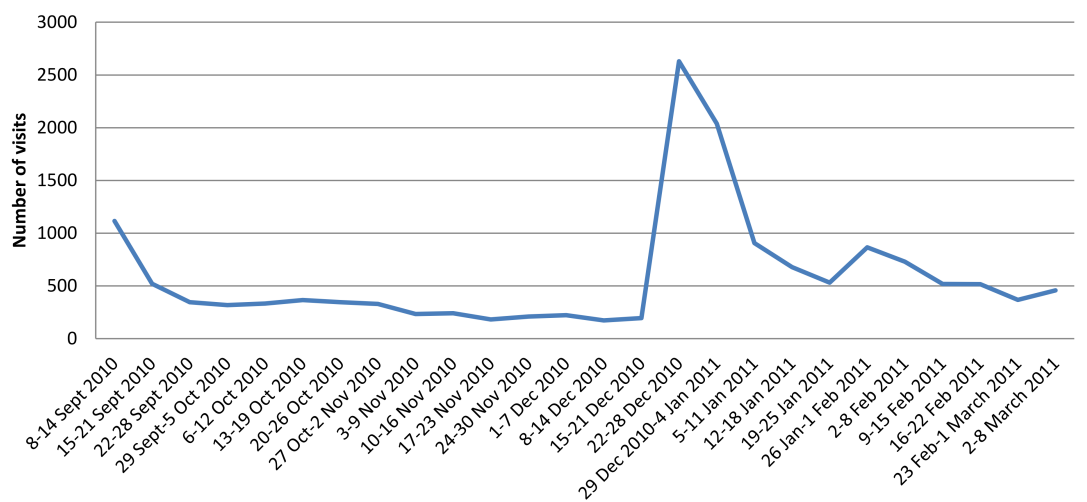


Figure 7. Visits to the Transcription Desk, 8 September 2010 to 8 March 2011

Following the publicity surrounding the launch of the Transcription Desk, there were 1,115 visits to the site during the first week, though things settled down soon afterwards when, during the remainder of Period One the site subsequently received an average of forty visits per day. Indeed, in November and December the number of daily visits rarely rose above thirty, on occasion reached sixty, but dropped as low as seven during mid-to-late December. Traffic to the Transcription Desk had essentially flatlined, though the volunteers then taking part had transcribed 350 manuscripts by the time Period One ended.

Then came the *NYT* article. From eleven visits on 26 December, traffic rocketed to 1,140 visits on 27 December, with a further 1,411 the following day. Remarkably, thirty per cent of all visits during the testing period to the Transcription Desk came between 27 December 2010 and 4 January 2011. The *NYT* article also had the effect of increasing the regular level of traffic to the site, to an average of 141 visits per day. The number of visits did not regularly drop below 100 until 19 January, and from then to 8 March the site rarely received fewer than sixty per day. In short, the publicity derived from the *NYT* article provided a level of traffic and an audience of potential volunteers which it is hard to see how we would have otherwise reached.^[24]

The Transcription Desk has been visited by users from ninety-one countries (Figure 8); most visits over the six months were from the United States, with the UK in second place.^[25] This again reflects the *NYT*'s impact and lack of comparable British press coverage during Period Two, as during Period One there were more than double the number of visits from Britain as there were from the United States (Table 1).

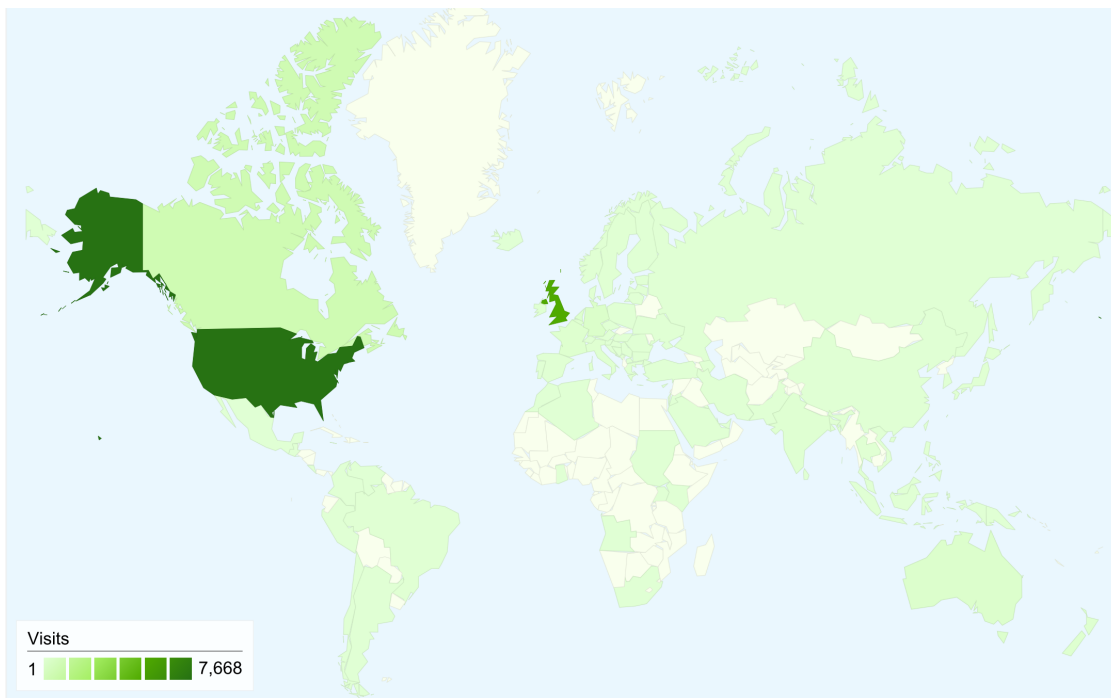


Figure 8. Google Analytics map showing from where the Transcription Desk was accessed, 8 Sept 2010 to 8 March 2010

Country	No. of visits
United States	7,668
United Kingdom	7,668
Canada	757
Germany	246
Australia	198
Netherlands	148
France	125
New Zealand	92
Belgium	86
Italy	71

Table 1. Countries from which the Transcription Desk was Most Accessed, 8 Sept 2010 to 8 March 2010

Volunteer Registrations

Figure 9 shows the number of registered volunteers, the number of manuscripts transcribed, and the number of transcripts signed off as complete. During Period One, twenty-three accounts were registered on average each week, though this dropped to eight per week during the weeks ending 26 November to 23 December 2010. Publication of the *NYT* article saw the volunteer base more than treble from 350 at the end of Period One, to 1,087 when the statistics were next recorded on 7 January 2011. However, after this flurry of activity, the weekly average of newly registered users dropped to twelve, lower even than during Period One. Indeed, from the week ending 18 February to 8 March, week-on-week growth of the overall user base dropped to less than one per cent, with an average of five new registrations per week.^[26]

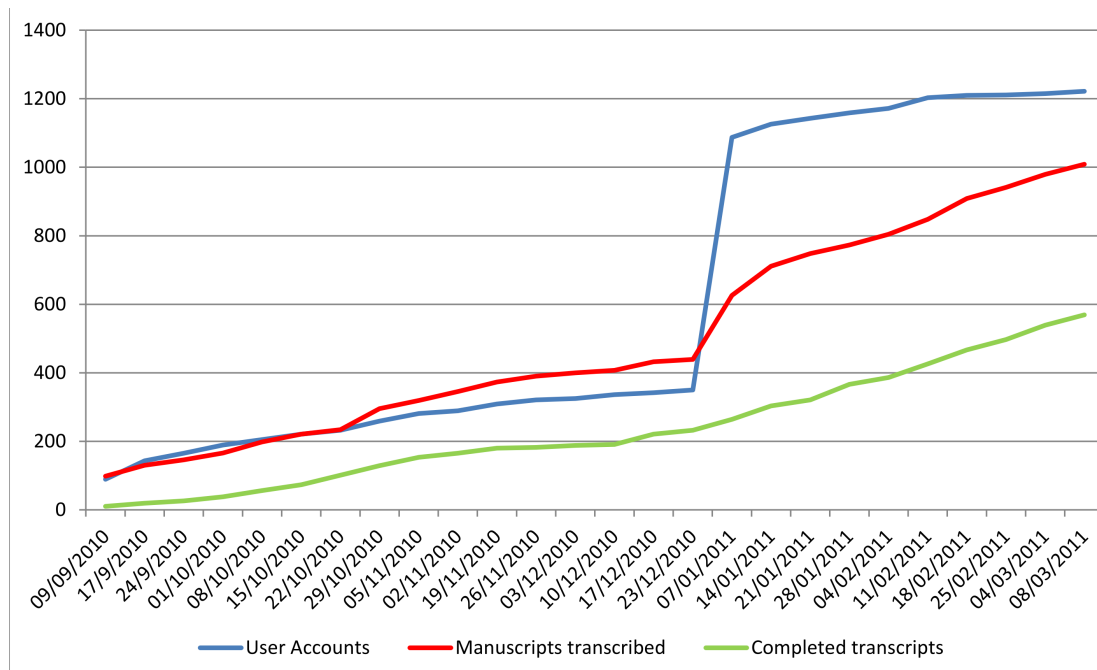


Figure 9. User accounts registered, manuscripts transcribed, and manuscripts locked, 8 September 2010 – 8 March 2011

How Were Volunteers Recruited?

Perhaps unsurprisingly given the project's nature, most respondents came to hear of *Transcribe Bentham* through online media, with the *NYT* article – which was published online as well as in print – being the single most commonly cited source (Figure 10). Our approaches to the British media were largely in vain, while the American press – including the *Boston Globe*, the *Chronicle of Higher Education*, as well as the *NYT* – reported on the progress of the project much more enthusiastically. We did not send out a press release in the United States or in Europe; the journalists who reported on *Transcribe Bentham* in the press and on the radio approached us for information, and in this regard, the blogosphere seems to have been an important means of attracting media attention. Other than sending out additional press releases in the UK, it is unclear what more we could have done to promote the project to the British media; it is also worth noting that historically, the Bentham Project has had trouble gaining media attention for Bentham studies. That said, the British press were not entirely unreceptive: an article on academic crowdsourcing, which included a report on *Transcribe Bentham*, appeared after the testing period, in the 11 September 2011 edition of *The Sunday Times* [Kinchen 2011], while another small piece appeared in the November 2010 edition of *Wired* for the iPad. It appears that it was traditional media in the United States – albeit their web presence – which was the most important recruiter of volunteers.

Google Adwords was a failure for us as a recruitment strategy. Our advert was displayed 648,995 times, resulting in 452 clicks, but sent no visitors to the Transcription Desk. The team's lack of experience in using Adwords may account for this failure; regardless, it was considered too costly an experiment to persevere. Social media such as Twitter and Facebook, while raising awareness of *Transcribe Bentham*, also appears to have had little impact in driving traffic directly to the site, despite staff using them on a routine basis for publicity, communicating with volunteers, and issuing notifications.^[27] Twitter, in particular, appears to have been more a means by which technology-savvy users have followed *Transcribe Bentham's* progress, rather than acting as a recruitment method. Nevertheless, social media has helped to integrate the community to an extent, by providing a platform for editors to share results and keep volunteers and other parties informed.

Time and budget constraints impacted upon the success of the publicity campaign. We spent around £800 of our budget, though the campaign cost more overall, as we used additional funds acquired through a UCL Public Engagement bursary and other sources. With more funds we might have been able to try different strategies, and a

commercial advert in a history magazine might have been possible. It also took time for the press release, distributed at the launch of the Desk in September, to yield results. Project staff liaised with journalists soon after the launch but, owing to the production process, stories often did not appear for some time after. *Wired* published its piece in November while *The New York Times* article appeared on December 27th – unfortunately when project staff were on holiday and unable fully to capitalise on the publicity or deal with enquiries immediately – some months after the launch. Thus, the user base only started to grow, and regular participants emerge, as the project's funded testing period moved into its final three months.

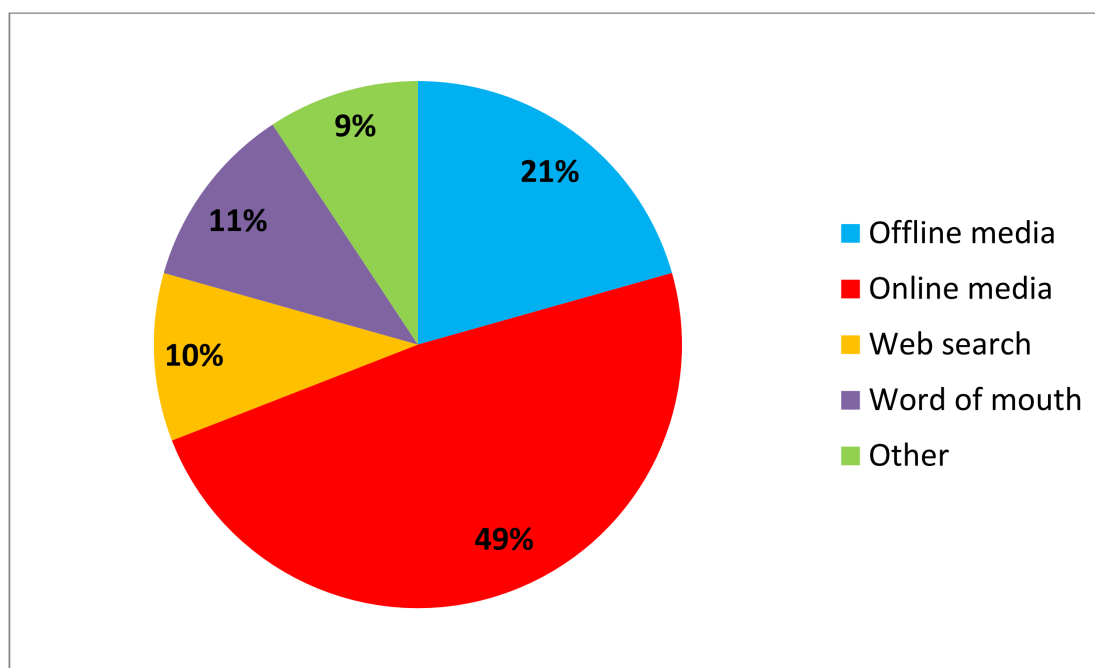


Figure 10. Means by which survey respondents first became aware of *Transcribe Bentham*

[28]

Who Were the Volunteers?

Analysis of user survey responses and user profile pages suggests that most of those interested in taking part in *Transcribe Bentham* were educated professionals (Figure 11). Ninety-seven per cent of survey respondents had been educated to at least undergraduate level, and almost a quarter achieved a doctorate. Almost two-thirds of survey respondents were female,^[29] whereas, for example, those taking part in *Galaxy Zoo* were overwhelmingly male [Romeo and Blaser 2011] [Raddick et al 2010].^[30]

39

Over a third of respondents were either academics or higher education students of varying levels, while those working in the arts, editing, and IT also registered with the project (Table 2). Almost a tenth of respondents were retirees and – from information gleaned from their user profiles – it appears that at least two of the most active transcribers were retired, while the most prolific volunteer had taken a year out from study. None of the regular volunteers were university students, and nor were any school pupils or teachers; the outreach strategy to London schools thus failed to generate sustained participation. Only one school responded to the letter and, while the Raines Foundation School showed enthusiasm at their visit to the Bentham Project, there was no sustained activity on the site. With an enhanced publicity campaign and more time, we would have been able to expand our outreach to schools by targeting independent schools, which may have had more resources available to dedicate to such a project.

40

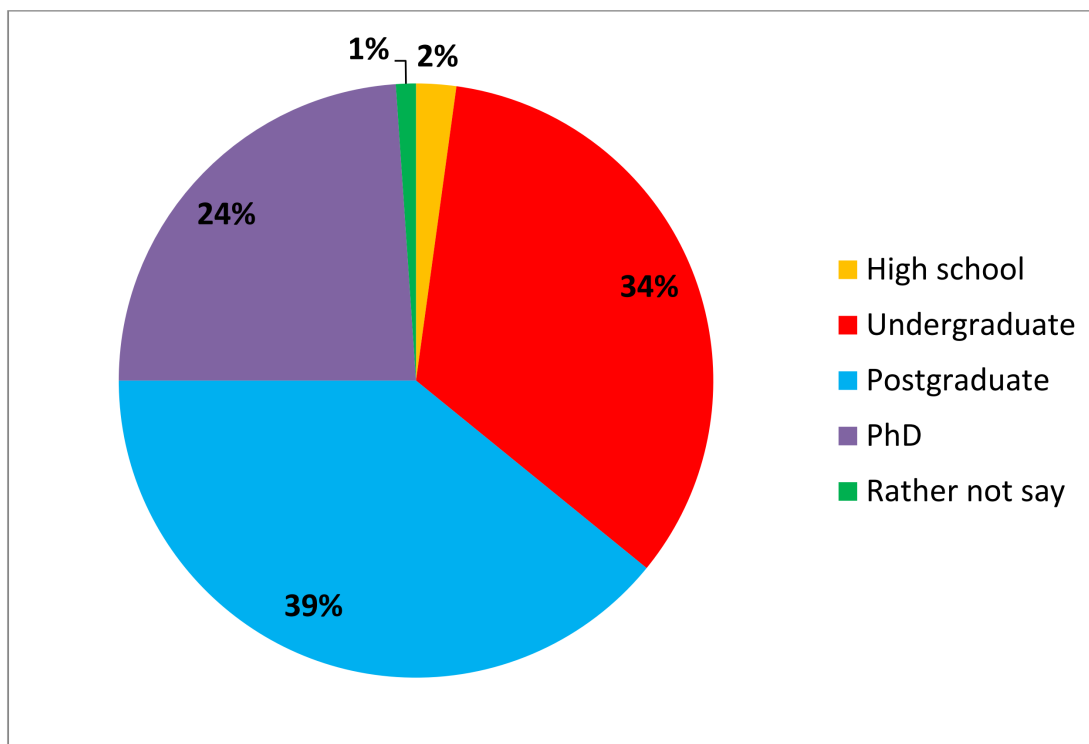


Figure 11. Education of *Transcribe Bentham* user survey respondents

[31]

Sector	No. of respondents (<i>percentage</i>)
Academia	14 (17)
Administration	5 (6)
Arts	8 (10)
Civil servant	3 (4)
Editing/publishing	8 (10)
Engineering	2 (2)
Finance	3 (4)
IT	4 (5)
Librarian	1 (1)
Project management	1 (1)
Psychiatrist	1 (1)
Researcher	3 (4)
Retail	1 (1)
Retired	7 (9)
Student	15 (18)
Teacher	1 (1)
Transcriber	2 (2)
Unemployed	3 (4)
Total	82 (100)

Table 2. Occupations of *Transcribe Bentham* user survey respondents

Holley found that while the age of crowdsourcing volunteers “varies widely”, the most active were mainly “a mixture of retired people and young dynamic high achieving professionals with full-time jobs” [Holley 2010]. Our results generally agree, though our survey results would suggest that *Transcribe Bentham* volunteers were, generally speaking, perhaps a little older. Most respondents were at least forty-one years of age and, notably, a fifth were over sixty. Where it is possible to glean information, the most regular transcribers appear to be in these upper age brackets. These results emphasise, at least for a project like *Transcribe Bentham*, the importance of recruiting users with plenty of disposable free time, and those associated with academic and/or professional backgrounds. Our publicity outreach to these audiences was, therefore, worthwhile.

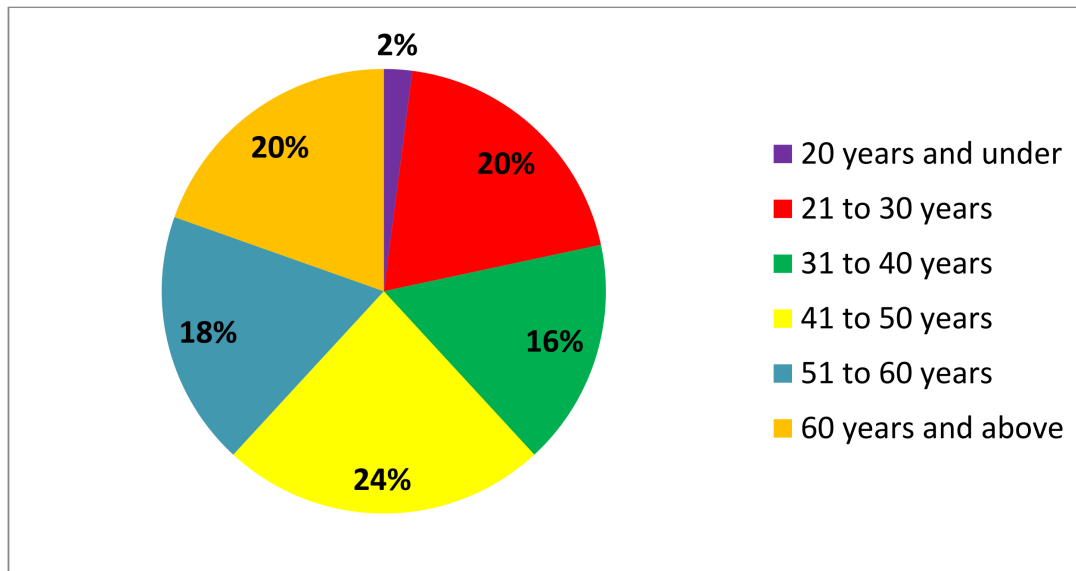


Figure 12. Age Profile of Transcribe Bentham User Survey Respondents

Only a minority of respondents had worked with manuscripts, or had any specialised training in reading historical handwriting, prior to taking part in *Transcribe Bentham*.^[32] There is thus little sense that *Transcribe Bentham*, in spite of publicity efforts to engage that circle, tapped into a pool of experienced and/or trained historical transcribers. However, a handful of respondents did have some prior experience, and they included a medievalist, a historical editor, and one particularly experienced user who had previously transcribed 25,000 documents dealing with the Canadian government’s relationship with Aboriginal Canadians.

What Was Produced?

Assuming that all 7,441 unique visitors to the Transcription Desk (as recorded by Google Analytics) were indeed individuals, then only a maximum of six per cent of those who visited the site were moved to register an account. Yet the volume of work done by relatively few was formidable indeed. 439 manuscripts were transcribed during Period One at an average rate of twenty-three per week, though only one volunteer then submitted work on a consistent basis. Even after the enthusiasm of September and October gave way, an average of eighteen manuscripts per week were transcribed during November and December, though admittedly parts of the latter month were very quiet.

Transcribe Bentham really began to fulfil its potential in Period Two as media attention paid dividends, though sadly just as the funding period was coming to a close. 187 manuscripts were transcribed between 27 December 2010 and 7 January 2011 alone, an increase of 43% on the end-of-Period One total. During Period Two, seven volunteers transcribed substantial amounts of material on a regular basis, producing an average of fifty-seven transcripts produced each week until 8 March.^[33]

Over the testing period as a whole, volunteers transcribed an average of thirty-five manuscripts each week^[34]; if this

rate was to be maintained then 1,820 transcripts would be produced in twelve months. Taking the complexity of the task into consideration, the volume of work carried out by *Transcribe Bentham* volunteers is quite remarkable (Table 3). Some manuscripts are only a few words long, while others can be up to two thousand words in length, and we estimate that the average manuscript is around 500 words long (plus mark-up). *Transcribe Bentham* volunteers transcribed over 500,000 words (plus mark-up) during the six-month testing period^[35]

Month	Total manuscripts transcribed	Average no. of manuscripts transcribed per week
8 Sept to 24 Sept 2010	68	23
25 Sept to 29 Oct 2010	149	30
30 Oct to 26 Nov 2010	95	24
27 Nov to 23 Dec 2010	49	12
24 Dec 2010 to 28 Jan 2011	334	84
29 Jan to 25 Feb 2011	168	42
26 Feb to 8 March 2011	68	n/a

Table 3. Number of Manuscripts Transcribed, 8 September 2010 – 8 March 2011

At the end of Period One, 53% of all transcripts were judged to be complete, but this dropped to 42% early in Period Two as large numbers of new users began to transcribe and grappled with Bentham's handwriting and style, and the transcription interface.^[36] The proportion of completed transcripts remained around this level until late January, when an "Incomplete Folios" category was made available to provide volunteers with the option of working on transcripts requiring improvement. By 8 March, 55% of transcripts were completed and locked.^[37]

Super Transcribers

Of the 1,207 registered volunteers, 259 (21% of the overall user base) transcribed manuscripts during the testing period (Table 4).^[38] Most worked on only one manuscript; this is not to say that they transcribed it to *completion* – although many did so – but that they transcribed at least some part of it.

No. of manuscripts transcribed	No. of volunteers (<i>percentage</i>)
1	163 (63)
2 to 5	74 (28)
6 to 9	7 (3)
6 to 9	7 (3)
10 to 19	7 (3)
20 to 30	1 (<)
63 +	7 (3)
Totals	259 (100)

Table 4. Number of Manuscripts Transcribed by Volunteers

A smaller number of volunteers contributed, however, the majority of the work. Fifteen volunteers transcribed between six and thirty manuscripts each, but only a further seven could be described as active on a regular basis, and thus analogous to Holley's "super volunteers", individuals which she found to comprise the highly-motivated backbone – usually around ten per cent – of a crowdsourcing project's user base, and who contribute the majority of the labour [Holley 2010]. Johan Oomen and Lora Aroyo also note that despite "an explosion of user-generated content on the Web, only a small portion of people contribute most of it". They estimate likewise that only ten per cent of online users

generate content, and that a mere one per cent of these generators “actively and consistently contribute the majority of the user-generated content” [Oomen 2011].

The seven volunteers shown in Table 5 comprise a mere 0.6 per cent of all registered *Transcribe Bentham* users – and three per cent of those who did transcribe – yet these “super transcribers” had, by 8 March 2011, between them worked on 709 (or 70%) of all of the 1,009 transcribed manuscripts.

49

	No. of manuscripts transcribed	Percent of total transcribed (1,009)
Volunteer A	80	7.9
Volunteer B	71	7
Volunteer C	280	27.8
Volunteer D	67	6.6
Volunteer E	66	6.5
Volunteer F	82	8.1
Volunteer G	63	6.2
Totals	709	70.1

Table 5. “Super Transcribers” and Their Work

[39]

50

Only Volunteer C had been with the project since around launch time, having begun taking part on 22 September 2010, and she worked on a staggering 28% of all transcribed manuscripts.^[40] All of the other early volunteers during Period One had either not returned to the project or transcribed only on an intermittent basis. The remaining six super transcribers were all recruited as a result of the *NYT* article and, as far as we can tell, were all based in North America.

Holley found that “super volunteers” can come to regard working on a project as a full-time job [Holley 2010]. While this is an exaggeration in the case of *Transcribe Bentham*, it is certainly the case that super transcribers spent significant amounts of time working on manuscripts.^[41] It is also clear that the super transcribers care a great deal about *Transcribe Bentham*, and they follow with huge interest announcements, improvements, and other developments related to the project. There is an evident feeling of responsibility on their part at being entrusted with the material, and they take great pains to ensure their work is accurately transcribed and encoded; indeed, submissions by super transcribers generally require minimal editorial intervention. They have also made a number of helpful suggestions which have been implemented, such as providing samples of Bentham’s handwriting and creating a page to list partially-transcribed manuscripts.

51

Transcribe Bentham did, therefore, recruit an active, though small, user community engaged in heavyweight peer production as well as a crowd of one-time users.

52

Sustaining Motivation

As Romeo and Blaser note, “[u]nderstanding the motivations of online volunteers is an important way to inspire and sustain participation” [Romeo and Blaser 2011]. Those responding to our survey reported that they were mostly motivated to take part by a general interest in history and philosophy, in crowdsourcing and the technology behind the project, and in Bentham himself (Figure 13). A significant number cited contributing to the greater good by opening up Bentham’s writings to a wider audience as a motivation, and several even found transcription fun! Though our survey sample is relatively small, similar findings by others examining participant motivations suggest that the results are sound. Peter Organisciak found that primary motivational factors included subject matter, ease of participation, and a sense of making a meaningful contribution, while the top five NLA text correctors cited their enjoyment of the task, interest in the material, recording Australian history for posterity, and providing a “service to the community” as their main motivations for taking part [Organisciak 2010, 83–88] [Holley 2009, 17–18].

53

Transcribe Bentham consciously implemented means to encourage volunteers through friendly competition, or what Romeo and Blaser describe as “game like mechanics” [Romeo and Blaser 2011]. The “Benthamometer” illustrated the project’s overall progress, and the points, multi-tiered ranking system, and leaderboard provided a visible recognition of work and progress. Other crowdsourcing projects have considered such features to be important, and it was thus surprising that *Transcribe Bentham* volunteers regarded “competition” and “recognition” to be of such low importance. Organisciak likewise found points systems, achievements, and leaderboards to be of secondary importance to volunteers [Organisciak 2010, 90].

54

However, this finding comes with a significant caveat: the most active super transcriber was indeed partly motivated by both competition and recognition, and it is clear that projects must be flexible enough to meet the motivational needs of a variety of volunteers.^[42] Perhaps others were subconsciously competitive, as another volunteer noted that “I’m usually a very noncompetitive person but [...] became inspired to do more when I find I’m losing my ‘position’” [on the leaderboard] – it’s ridiculous! but quite fun”.^[43]

55

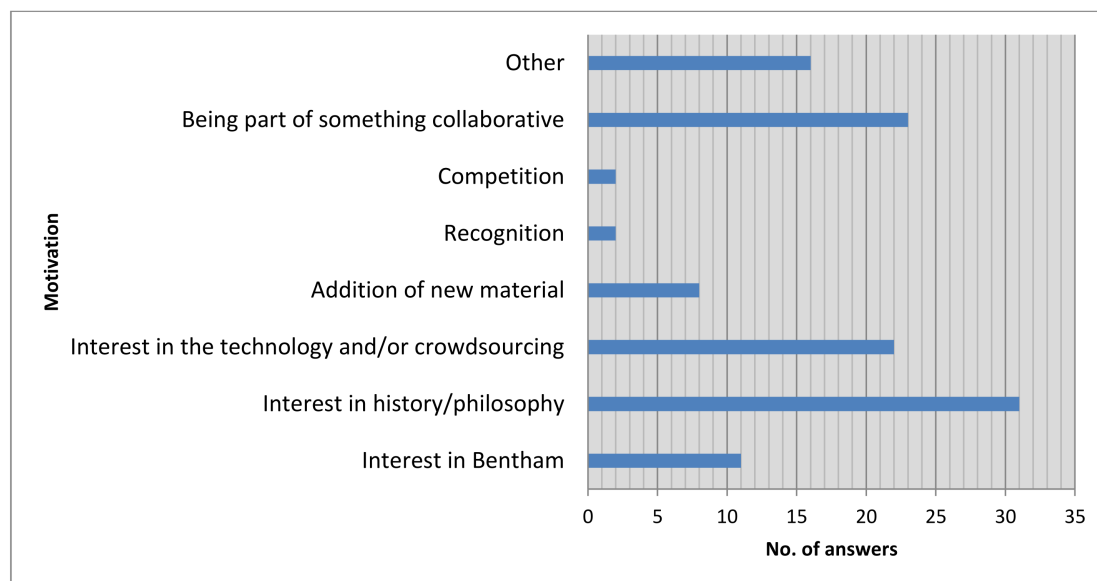


Figure 13. Survey respondents' motivations for taking part in *Transcribe Bentham*

[44]

56

Many respondents found the intellectual challenge and puzzle-solving aspect of transcription enjoyable and highly rewarding, with several comparing deciphering Bentham’s handwriting to solving a crossword or sudoku. Two transcribers were excited by working on manuscripts which may not have been looked at since Bentham wrote them, while contributing to the accessibility of the digital collection and the production of the *Collected Works* was also a factor for some.^[45] One volunteer notably described *Transcribe Bentham* as a “literary form of archaeology”: “Instead of using a brush to uncover an object, you get to uncover historical information by reading and transcribing it. It leaves his legacy available for all to access.”

This is an important point, and supports Holley’s suggestion that volunteers are more likely to take part in projects run by non-profit making organisations, as opposed to those run by commercial companies [Holley 2010]. When volunteers submit transcripts to *Transcribe Bentham*, they agree to the following disclaimer: “You are contributing to an Open Access initiative. For the purpose of UK legislation, copyright for all materials within *Transcribe Bentham*: Transcription Desk, including this transcription, resides with UCL (University College London). If you do not want your writing to be edited and redistributed at will, then do not submit it here.”

57

Overwhelmingly, survey respondents had no problem with this, which speaks to the mutual trust and respect between a project and its volunteers which is vital for success. Users recognised that UCL had made freely available priceless

58

manuscript material, a transcription interface and instructions, and that their work was a transcript of Bentham's writing rather than original intellectual content. Such approval did, however, come with the proviso that the submitted transcripts would not be commercially exploited, and would remain accessible to all in the future; the free availability and long-term preservation of the transcripts in UCL's digital repository certainly meets these requirements.

The vast majority of those who visited the Transcription Desk did not become active users, and it is thus just as important to understand what dissuaded people from participating as it is to understand what encouraged them (Figure 14). Some were daunted by the task, or felt that their contributions might not be of sufficient standard to be deemed worthy. For example, one respondent noted that he had been unable to read Bentham's handwriting "with enough clarity to be happy recording my guesses in an important academic project," while another "confess[ed] to being a bit intimidated [...] fear of failure?" This is a shame, as any contribution to *Transcribe Bentham* is beneficial to the project; perhaps in the future the project team could provide more reassurance to volunteers that all efforts are valuable. Another respondent decided that she could not take part because "I need a paying transcription position"; *Transcribe Bentham* is in no position to offer remuneration, and – as discussed above – it was considered that offering money for participation would be contrary to both the project's collaborative principles and spirit.^[46]

It is notable that no survey respondents mentioned the importance of feedback from project staff or prompt responses to submissions. However, we have discovered through correspondence and observation that volunteers greatly appreciate notifications as to whether or not their submissions have been accepted, as well as responses to email requests for further information. Indeed, volunteer enthusiasm can noticeably fall away when feedback and acknowledgement are not given, and we may have lost potential long-term contributors during the December break when staff were away. This suggests that project staff must devote time to answering queries from users and to maintaining social media features in order to facilitate interaction, otherwise the project can appear dormant, and volunteers may lose interest or feel disconnected and exploited.

The single most important dissuading factor, however, for those who responded to the survey, was time, or rather the lack of it, in which to transcribe and encode, or learn how to do so. Reading and deciphering the manuscripts was found to be the main challenge: as one respondent replied, "[t]he real difficulty is in reading [Bentham's] handwriting". Indeed, over half of respondents found that deciphering Bentham's hand took longer than encoding. Though text-encoding was an additional complication to the process, encouragingly few survey respondents found it prohibitively difficult.^[47] However, this comes with the significant qualification that almost two-thirds of respondents found that it took some time to get used to the encoding process, and that a significant proportion thought they spent as much time encoding as they did deciphering Bentham's handwriting.^[48]

Indeed, most free-text responses noted that encoding (and the extensive instructions on how to do so) were daunting to beginners, but that the process became reasonably straightforward with practice.^[49] Several respondents were not so sure: one believed there was "[t]oo much markup expected", another that encoding was "unnecessarily complicated", and one – who, unsurprisingly, is not a regular transcriber – found encoding "a hopeless nightmare" and the transcription process "a horror". To many, text encoding was simply a time-consuming complication and almost certainly dissuaded participation. As one respondent put it, "The transcription process has been a little more complicated than I anticipated. I thought I would be able to type away but I have to learn a little more before I get started, so I'm not wasting anyone else's [sic] time."

Some proved very able at text-encoding, to the point where almost half of survey respondents were confident enough to type at least some tags manually rather than use the toolbar. However, the negative views of encoding are very real concerns which must be addressed when improving and refining the transcription tool.^[50] Indeed, the majority of survey respondents were unfamiliar with the principles and practice of text encoding before taking part in *Transcribe Bentham*, though with a few exceptions: some were aware of the basics through editing *Wikipedia*, writing HTML, or using early word processors. One volunteer in particular noted, that "I'm part of the wiki generation -- we have encoding encoded into our blood".^[51] While crowdsourcing projects should not underestimate the IT proficiency of their volunteers, or their ability to learn new skills, the task at hand should be made as simple as possible.

As such, we made the transcription and encoding process, and the instructions, as user-friendly as we could given our requirements and our short time-frame. However, these survey responses and the project's inability to retain all but a few long-term volunteers suggests that the task appeared too complex and time-consuming for many prospective volunteers. Though we can offer guidance and assistance to those attempting to decipher Bentham's handwriting, evidently, simplification of the transcription process and instructions would improve user recruitment and retention. For example, the introduction of a What-You-See-Is-What-You-Get interface, as an alternative to the transcription toolbar, with encoding occurring "under the hood", would prove an attractive option for many volunteers. They would then be able simply to concentrate on transcription without having to digest lengthy instructions [Causer, Tonra and Wallace 2012].

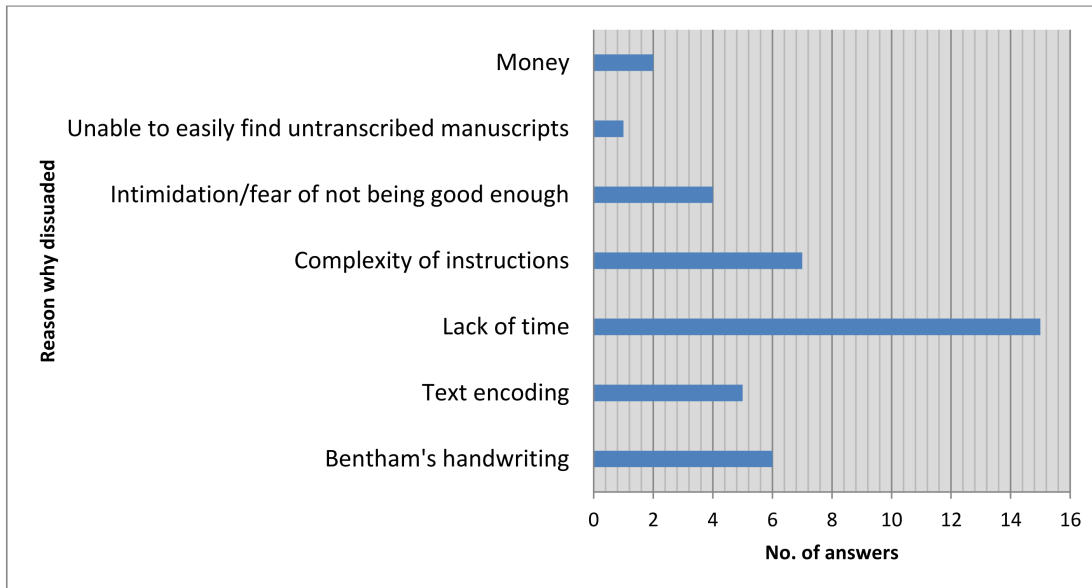


Figure 14. Factors which dissuaded survey respondents' participation in *Transcribe Bentham*

[52]

User Interaction and Community Cohesion

Fostering a sense of community would, we hoped, stimulate participation by encouraging users to strive to achieve common goals. Our community building strategies, however, appear to have had limited success. For example, only one pre-existing transcriber attended the public outreach events held in May 2011. This transcriber did appreciate the opportunity to meet project staff and experience face-to-face interaction. Of course, the volunteers who live outside of the UK were inhibited from attending these events and thus were unable to integrate physically with the community. In terms of fusing the *Transcribe Bentham* community together therefore, this particular programme of public events was ineffective.

65

Furthermore, despite the project having integrated social media facilities into the Transcription Desk, there is minimal evidence of interaction between *Transcribe Bentham* users. We are only aware of a handful of occasions on which users have communicated publicly with each other via their user page, a transcript, the discussion forum, or the project's Facebook page. While users may have exchanged private messages or emails, the survey results suggest otherwise: only one respondent reported that she had contacted another via their user page, two had done so via the discussion forum, and another via Twitter.

66

This is surprising, as it was anticipated that users would demand such social functionality. Indeed, the *North American Bird Phenology Project's* survey respondents specifically requested a discussion forum to communicate amongst themselves, and the NLA was met with requests from text correctors to provide a facility by which users could contact each other [Phenology Survey 2010] [Holley 2009, 23]. Many of the major crowdsourcing projects have discussion

67

forums in which volunteers are active to one degree or another.^[53] These are certainly busier than the *Transcribe Bentham* forum which – rather than being a space for volunteers – became merely another avenue by which users contact project staff or in which project announcements are made. Perhaps a larger volunteer community is required for the forum to be used to a greater extent.

Moreover, by the end of the testing period, seventy-eight per cent of survey respondents stated that they had not added any information about themselves to their user profiles.^[54] A cursory look through registered users' pages – even those belonging to super transcribers – reinforces the impression that volunteers are not particularly interested in social media facilities, or perhaps even being part of an online community *per se*, as hardly any contain any information beyond a name and, in a few cases, a photograph or avatar. Survey responses suggest why this may be the case. One respondent had “added minimal information because I don't know enough about the community [...] to want to make my information available to others”. Others were concerned that such information “seems unnecessary”, that they were “[n]ot sure how this would be helpful”, thought personal information would be “more interesting and useful to the T[ranscribe]B[entham] organization than to individual contributors”, or didn't see the feature as important. Two further respondents expressed doubts about the wisdom of putting too much “personal profile info on public sites”, which may be a significant issue for future crowdsourcing projects seeking to create a community. Furthermore, while there have been technical issues with the site's social features, transcribers do not appear too concerned about these problems as we have received no enquiries about their non-functionality.

On the other hand, some wanted to know more about their fellow volunteers. One believed that “it is important for each person involved in the project to at least let us know their educational background and philosophy. It encourages discussion”. Another thought that it “is only possible to create an online community if people are willing to say something about themselves”. However, they were in a definite minority, with most seemingly uninterested in, or unaware of, the site's social media capabilities.

Transcribers have proven wholly good-natured in going about their work: there have been a couple of easily-resolved but unintended revisions of others' work, but certainly no evidence of “edit warring”, repeated reversions, or arguments. There is certainly none of the overt territoriality exhibited in the tagging of text in one NLA newspaper, when an anonymous user requested a registered text corrector to “please refrain from appropriating and/or inserting any [text] correction or tag” concerning a particular historical figure, whom the anonymous user appeared to believe belonged to him or her.^[55]

That said, there is no evidence that *Transcribe Bentham*'s volunteers have collaborated directly upon manuscripts; rather, they participate in what might be termed unconscious collaboration, with volunteers sometimes adding another layer of work onto what went before. As one respondent to the survey suggested, it “[s]eems to me the whole system is a collaboration, since we're encouraged to submit unfinished transcriptions & others can work on them. Maybe a serial collaboration”. Two respondents were unaware that direct collaboration was possible, while another did not “feel secure enough with coding to muck up someone else's work”, and presumably worked instead on unedited folios.

Indeed, three-quarters of edited manuscripts were worked upon by only one volunteer, while relatively few manuscripts were transcribed by three or more. Those manuscripts transcribed by multiple volunteers were generally begun with by one or more new users and otherwise “abandoned”, and later completed by a more experienced transcriber. This all suggests that volunteers appeared to prefer starting transcripts from scratch, and to work alone (Table 6), with communication and acknowledgement from staff being of much greater importance than collaboration with other users. This might appear counter-intuitive, as dabbling with partially-completed transcripts could allow volunteers to get used to the project and transcription process, and be less daunting than starting a transcript from scratch. The advent of the incomplete folios category in mid-January 2011 did see regular volunteers finishing partially-complete transcripts, though this option has not been popular with new volunteers, who seem to prefer to start with an untranscribed manuscript. Indeed, we have regularly received complaints from volunteers about being unable easily to distinguish between untranscribed, partially transcribed, and completed transcripts. Perhaps registration acts as a barrier to those who might wish to tweak a little bit of encoding or transcription here and there.

No. of transcribers	No. of manuscripts (<i>percentage</i>)
1	753 (74.6)
2	256 (25.4)
2	753 (74.6)
3	64 (6.3)
4	16 (1.6)
5	5 (0.5)
6 or more	3 (0.3)
Total	1009 (100)

Table 6. How Many Volunteers Transcribed Each Manuscript?

4. Conclusion

Through the intensive publicity campaign discussed above, the *Transcribe Bentham* team has successfully promoted its project, and crowdsourcing and manuscript transcription more generally, to a wide audience. Our findings may offer suggestions to those who plan to build volunteer communities for their own projects, and specifically for those looking to crowdsource more complex tasks like manuscript transcription rather than simple data entry or OCR correction. Some of our conclusions are derived from our volunteer survey; though the sample was relatively small, its findings are reinforced by analysis of volunteer transcriber profiles, statistical data from Google Analytics, and general observation of volunteer behaviour. Furthermore, that the survey results also fit into the broader narrative of discussions of crowdsourcing participants' characteristics and motivations, also suggests that our findings are sound.

We found that *Transcribe Bentham* volunteers were mostly well educated, often associated with academia or had a professional background, and had a prior interest in history and philosophy; many of the crowd also had an interest in digital humanities and were IT literate, but few were associated with schools. This highlights the importance of targeting the correct audience for the task at hand, in this instance, transcribing historical material. Most of the transcribers were recruited through online media, particularly the article in the *New York Times*, which emphasises the importance of securing the attention of a major news organisation with both national and international circulation. Our press release was certainly worthwhile, therefore, though unfortunately the British press displayed a disappointing lack of interest in the story.^[56]

The crowd of users were generally engaged in lightweight peer production. Most contributed only once to the site and were most likely put off from further participation, our results suggest, by a lack of time, which, unfortunately, there is little we can do to solve. However, this was exacerbated by the complexity and difficulties involved in transcribing or encoding. The processes involved in such an endeavour need to be simplified as far as possible to ensure increased and ongoing volunteer recruitment and retention, and to make the most of their available time. Indeed, wasting the time of users is perhaps the worst sin any crowdsourcing project can make.

Most of the *Transcribe Bentham* crowd failed to interact with each other, to use the social features of the site, or to respond to the stimulation we hoped would be provided by the Benthamometer and the leaderboard. Improved social features might facilitate interaction, though the survey results and observation of the site suggest they might not be that important to some volunteers. Direct collaboration – if thought desirable – could be further facilitated by improving the interface's functionality, or by providing further instruction. Respondents to the survey suggested that we provide a way for the inexperienced to request help from other users – an experienced users' contact list, for example. Giving experienced and motivated volunteers moderator status may be one way in which crowdsourcing projects could improve community cohesiveness, and is something we would like to explore in the future.

Unfortunately, owing to *Transcribe Bentham's* short time-frame due to the limited funding budget, the project was unable properly to benefit from many of its recruitment strategies; the impact of international press attention was only starting to

be felt as the six-month testing period was ending. Too heavy reliance upon the media for publicising a time-limited project is a risk, as the timing of media attention is to a large extent beyond a project's control owing to changing media deadlines and priorities. If a piece comes out too soon, a project may not yet be available to the public and the attention cannot be capitalized upon; too late, and volunteers brought to the project as a result may be alienated if a cessation of funding causes the end of user support.

Transcribe Bentham, for example, was gaining momentum just as it was obliged to scale back owing to the nature of the AHRC grant; editors were unable to provide the detailed feedback to volunteers as they had done previously. As the entire project is built on mutual trust between project staff and the transcribers, we felt duty-bound to inform volunteers that the fully-staffed testing period would end on 8 March 2011. Several volunteers who joined the project as a consequence of the *NYT* publicity were hugely disappointed to learn of this only three months after they had first become aware of the project's existence, and had invested considerable time and effort in taking part. A number, quite understandably, stopped taking part altogether. Our crowd of 1,207 users was not transformed, therefore, into a cohesive and active community.^[57] These results suggest that projects requiring a volunteer community must plan their publicity strategies effectively and, if they have only a short fully-staffed and fully-funded live testing period, well in advance and, where possible, in conversation with interested media. Advanced planning might not be possible, of course, if, like *Transcribe Bentham*, the fully-staffed development stage is relatively brief, or the project is funded through a scheme issued at short notice.^[58]

Though it was a little disappointing that we failed to transform most of our crowd into regular volunteers, this is not necessarily a disadvantage for a community-based project. Various crowdsourcing projects have found that only a minority of users carry out most of the work, and so it proved with *Transcribe Bentham*; a small but remarkably active core of super transcribers became engaged in heavyweight peer production, providing a large quantity of transcripts. Community projects engaged in similar work must, however, be prepared and have the means to support this core group. Our super transcribers appreciated communication with and advice from staff, occasionally interacted with other users on the Transcription Desk and via social media, though few have populated their social profiles. One super transcriber attended the public outreach events held in May and evidently appreciated the opportunity to meet project staff in person, and at least one other responded positively to the multi-tiered progress system. This suggests, perhaps, that "team-building" features and strategies, in spite of their lack of appeal to some, should not be dispensed with completely.

Transcribe Bentham's volunteers produced 1,009 transcripts – or an estimated 250,000 to 750,000 words, plus mark-up – during the six-month testing period. On the face of it, this pales in comparison with the over 650,000 pages of ships' logs transcribed by *Old Weather* volunteers, the 36.5 million lines of OCR text corrected by volunteers involved in the NLA digitised newspaper collection, or the tens of millions of index entries keyed-in for Ancestry.com's *World Archives Project*. Yet, transcribing Bentham is more demanding than the often mechanical and repetitive tasks involved in other projects, and arguably lacks the more immediate popular and media appeal of citizen science and genealogy.^[59]

Building a volunteer community to carry out complex tasks can be beneficial and rewarding to humanities scholars, particularly at a time when they are under pressure to ensure that their work has impact. Nevertheless crowdsourcing does not necessarily render research projects more economical. Aside from potential significant short-term costs in digitisation, creating and testing a transcription tool, and promoting the endeavour, some long-term investment will also be required. Time and money will need to be spent on interacting with volunteers, maintaining and developing the transcription interface in response to volunteer needs, continual promotion of the project, and checking and offering feedback on submitted work. The experience of running *Transcribe Bentham* suggests that it is only over time that building a volunteer community pays off; though *Transcribe Bentham* by no means enjoys mass participation, those taking part currently produce transcripts at the same rate as had the Bentham Project employed a full-time transcriber [Causer, Tonra and Wallace 2012]. *Transcribe Bentham* has also raised awareness of the Bentham Project and engaged the public with its research to a greater extent than it has ever done before; volunteer community-based initiatives thus have the potential to enhance the profile and broaden the scope of research projects.

Transcribe Bentham's experience, then, demonstrates that there is an audience of potential volunteers who are willing

78

79

80

81

82

and able to engage in more demanding crowdsourced tasks, such as transcribing complex manuscripts. With a sustained and targeted publicity programme aimed particularly at educated retirees, students, and the relatively computer literate, an intuitive transcription tool, a task made as straightforward as possible, and staff committed to providing feedback and stimulating interaction, crowdsourcing the transcription of manuscripts – even those as challenging as Bentham’s – can be a successful and rewarding venture for volunteers and researchers alike.

Acknowledgements

We would like to thank our *Transcribe Bentham* colleagues: Philip Schofield, Justin Tonra, Richard Davis and his team, Martin Moyle, Tony Slade, and Melissa Terras; and to all of the volunteer transcribers who have generously given their time and effort, ensuring *Transcribe Bentham*'s ongoing success. We are grateful to Oliver Harris, Michael Quinn, Melissa Terras, Justin Tonra and Catherine Pease-Watkin for reading earlier drafts of this paper, and offering many very helpful comments. We would also like to thank our three anonymous reviewers, whose comments and suggestions have been invaluable in finalising the article.

83

Funding

This work was supported by the Arts and Humanities Research Council's Digital Equipment and Database Enhancement for Impact (DEDEFI) scheme, under which *Transcribe Bentham* was established. Please visit <http://www.ahrc.ac.uk/FundedResearch/Pages/ResearchDetail.aspx?id=149682> for further details.

84

Notes

[1] For a detailed cost-benefit analysis of *Transcribe Bentham*, see [Causer, Tonra and Wallace 2012].

[2] The British Library holds another 15,000 or so folios of material by and relating to Bentham.

[3] For example, the British government, in late 2010, announced plans to measure the nation's "happiness", with the first results due in 2012.

[4] For development of the transcription tool and UCL's digital repository, see [Causer, Tonra and Wallace 2012], and [Moyle 2011].

[5] The call for the AHRC's DEDEFI scheme was issued in September 2009, and the results were announced in February 2010. Interviews for the Research Associate positions were held in March 2010.

[6] Bentham's will, made shortly before his death on 6 June 1832, stipulated that his body be donated to medical science, and his skeleton be dressed and preserved as an "Auto-Icon". The Auto-Icon and its cabinet were brought to UCL in 1850, and it is on public display at UCL. See <http://www.ucl.ac.uk/Bentham-Project/who/autoicon> for more information, and http://www.ucl.ac.uk/Bentham-Project/who/autoicon/Virtual_Auto_Icon for the new "Virtual Auto-Icon", in which Bentham can be explored in fine detail.

[7] *Transcribe Bentham* video: <http://www.youtube.com/watch?v=CtEqW4WwMHU>.

[8] Full and up-to-date list of talks given by *Transcribe Bentham* team members (including in some instances audio and video): <http://www.ucl.ac.uk/transcribe-bentham/talks/>.

[9] Assistance on reading historical manuscripts: <http://www.ucl.ac.uk/transcribe-bentham/palaeography/>.

[10] TEI By Example: <http://tbe.kantl.be/TBE/examples/TBED06v00.htm>.

[11] Information relevant to A-Level and Higher study: <http://www.ucl.ac.uk/transcribe-bentham/information-for-schools/a-levels-and-scottish-highers/>.

[12] See the list of winners of the 2011 Prix Ars Electronica, <http://new.aec.at/prix/en/gewinner/2011/>, and the *Transcribe Bentham* citation, <http://prix.aec.at/winner/403/>. There were 407 entries in the Digital Communities category. The Award of Distinction thus constitutes joint second place, and carries a prize of €5,000. Pictures from the winners' forum are available (<http://www.flickr.com/photos/arselectronica/sets/72157627336428579/>), along with a video recording of the session (<http://www.youtube.com/watch?v=vvD3ipgZCTQ>).

[13] Digital Heritage Award 2011: <http://www.dish2011.nl/news/digital-heritage-award-2011-nominees>.

[14] The “Benthamometer”: <http://www.transcribe-bentham.da.ulcc.ac.uk/td/Benthamometer>.

[15] *Transcribe Bentham* leaderboard: http://www.transcribe-bentham.da.ulcc.ac.uk/td/Top_Users.

[16] *Transcribe Bentham* volunteer ranks: http://www.transcribe-bentham.da.ulcc.ac.uk/td/Help:User_levels.

[17] *Bentham in the Community* public events: http://www.ucl.ac.uk/Bentham-Project/news/events/events_archive/BITC.

[18] Spam was not a significant issue, and nor have any transcripts been vandalised with obscene language or content. The handful of attempts to spam the site saw the creation of user pages containing commercial links, or placing these links in the discussion forum or Facebook page. These were quickly deleted and the offending users blocked. Spam was often brought to staff’s attention by users. During July and August 2012, a number of spambot accounts were generated, though these caused no problems to the site and were easily removed.

[19] Progress updates are issued each Friday, via the discussion forum, project blog, Facebook, and Twitter accounts.

[20] The AHRC grant provided for the commission of a server log analysis, to be carried out by UCL’s Department of Information Studies. Owing to staffing issues within DIS, the log analysis was not completed, and we have instead relied upon Google Analytics for site statistics.

[21] The user survey was available from 25 January to 24 March 2011, and consisted of a combination of multiple-choice questions and free-text boxes in which respondents could add specific details and observations. It was open to all, and advertised via the project blog, *Facebook* page, *Twitter* feed, a site notice on the front page of the Transcription Desk, and by placing a message on each registered volunteer’s user page. A full copy of the questions asked is available from the authors upon request.

[22] To listen to the Deutsche Welle World radio feature, visit here: http://www.dw-world.de/popups/popup_single_mediaplayer/0,,14808024_start_0_end_0_type_audio_struct_3126_contentId_6424149,00.html. Dr Causer and Dr Wallace were interviewed for Austrian ORF Radio’s “Create Your World” series, leading up to the Ars Electronica festival. This can listened to here (in German): <http://oe1.orf.at/programm/280040>. Dr Causer was also interviewed in Linz by ORF Radio’s “Matrix” programme, which was broadcast on 29 January 2012 (<http://oe1.orf.at/programm/294290>).

[23] Visits from *Transcribe Bentham* staff offices were discounted from these figures. As of 3 August 2012, the site has received 34,372 visits from 14,008 unique visitors.

[24] From 9 March to 3 August 2012, the site has received a mean average of 66 visits per day.

[25] As of 3 August 2012, the Transcription Desk has been accessed by visitors from 115 countries.

[26] This pattern has continued since the end of the testing period. From 9 March to 3 August 2012, 487 new users have registered an account, at an average of seven new registrants per week. Apart from the periods 25 March to 1 April 2011, 9 to 16 September 2011, and 5 to 11 May 2012, week-on-week growth of the user base remains at less than 1 %.

[27] Of the 15,354 visits to the Transcription Desk between 8 September 2010 and 8 March 2011, the top five referring sites were: the *Transcribe Bentham* project blog (9,556, or 62% of all visits); Facebook (468, or 3%), Windows Live Mail (383, or 2.5%); Twitter (136, or 0.9%); and the *ArchivesNext* blog (91, or 0.6%). That most of the referrals were via the project blog emphasises that this was the main entry point for visitors. There is a corresponding dramatic increase in referral traffic from the project blog at the time of publication of the *NYT* article, which carried a link to the blog rather than the Transcription Desk. From 9 March 2011 to 3 August 2012, Facebook’s influence grew and drove around 20% of the traffic to the Transcription Desk during this period. Only 8% of referrals from Facebook, however, were from new users, which suggests that regular volunteers may have used Facebook as an entry point to the site.

[28] 92 respondents.

[29] There were 94 responses to a question asking “What gender are you?”. 54 (57%) said “female”, 36 (38%) said “male”, and 4 (4%) responded “would rather not say”.

[30] As of 3 August 2012, of the eight volunteers transcribing on a regular basis, seven are female.

[31] 92 respondents.

[32] There were 85 responses to a question asking “Have you had any palaeography training, or have you transcribed manuscripts before taking part in *Transcribe Bentham*?” 28 (33%) said “yes”, and 57 (67%) said “no”.

[33] This mean average for Period 2 is distorted somewhat by the elevated rate of transcription between 27 December 2010 and 7 January 2011. The median average for Period 2 is 43.

[34] The median average for the testing period was 27, and the modal average was 32. As of 27 January 2012, the overall mean average rate of transcription remains at 35 per week.

[35] On the same estimate, as of 3 August 2012, around 2 million words have been transcribed, plus mark-up. It is thus estimated that the 40,000 folios of material which were untranscribed at the start of the project contain around 30 million words. From 9 March 2011 to 3 August 2012, the mean average was 41 transcripts per week, though from 31 December 2011 to 3 August 2012 the rate was 51 transcripts per week.

[36] For more on the *Transcribe Bentham* moderation process, see [Causer, Tonra and Wallace 2012].

[37] Incomplete Folios category, http://www.transcribe-bentham.da.ulcc.ac.uk/td/Category:Incomplete_Folios. As of 3 August 2012, the proportion of completed transcripts has risen to 94%. This is mainly due to the increased proficiency of regular transcribers, as well as project staff working through incomplete manuscripts following the testing period.

[38] As of 3 August 2012, of the 1,726 registered volunteers, 311 (18%) have carried out any transcription.

[39] All usernames have been anonymised.

[40] As of 3 August 2012, Volunteer E has transcribed over 1,142 manuscripts, comprising around 28% of the total 4,014 transcripts. Volunteer D has, as of the same date, produced 709 transcripts (17%). Despite losing all but three of the then “super transcribers” at the end of the testing period, as of 3 August 2012, *Transcribe Bentham* now has eight volunteers transcribing on a regular basis. One user began transcribing in mid-September 2011, and has already transcribed 934 (23%) of the total, while another has produced 285 (7%) transcripts since 11 November 2011.

[41] There were 43 responses to a question asking “How long, on average, do you think you spend on *Transcribe Bentham* each week?” 32 (75%) spent up to 2 hours on the site per week, another 6 (14%) spent up to 6 hours, 2 (5%) spent up to 8 hours, and another 3 (7%) spent over 8 hours per week on the site. Unsurprisingly, the super-transcribers spent the most time per week transcribing.

[42] Organisciak notes that points systems are best suited “to serve the more dedicated contributors, giving them something to show for their dedication and something still to strive for” [Organisciak 2010, 90].

[43] Perhaps competition might be encouraged in the future by promoting short term goals, such as getting a set of manuscripts transcribed by a certain date.

[44] 43 respondents.

[45] All volunteers who have transcribed manuscripts will be acknowledged in the relevant published volumes of the *Collected Works*.

[46] [Organisciak 2010, 83] discovered financial reward to be a primary motivator in those projects which offer remuneration.

[47] There were 44 responses to a question asking “Have you found it easy to encode your transcripts?”. 13 (30%) said “yes”, 27 (28%) said “yes, once I got used to it”, and 4 (9%) said “no”.

[48] There were 43 responses to a question asking “Do you think you spend more time transcribing the manuscripts, or encoding your transcripts?” 24 (56%) said “transcribing”, 2 (5%) said “encoding”, and 17 (40%) said “an equal amount of time on both”.

[49] This result is most likely skewed, as those who were put off by the encoding in the first place were unlikely to complete the survey, or even to have been aware of it.

[50] There were 44 responses to a question asking “Do you use the toolbar to encode your transcripts, or do you type the tags manually?” 20 (45%) said they used the toolbar, 4 (9%) typed them, and 20 (45%) did both.

[51] There were 84 responses to a question asking “Were you familiar with the principles of text encoding prior to taking part in *Transcribe Bentham*?” 35 (42%) respondents said “yes”, and 49 (58%) said “no”.

[52] There were 36 responses to a question asking “If you do not consider yourself a regular user of *Transcribe Bentham*, or you have signed up but not transcribed, could you let us know what has put you off or dissuaded you from transcribing”?

[53] For example, see the NLA's *Trove* Forum (Digitised Newspapers sub-section) <http://trove.nla.gov.au/forum/forumdisplay.php?3-Digitised-newspapers-and-more>, accessed 13 April 2011, or *Galaxy Zoo* forum, <http://www.galaxyzooforum.org/>, accessed 13 April 2011.

[54] Several registered users filled in their social profile page in great detail, but did not transcribe anything; conversely, most super-transcribers have added little or no information to their profiles.

[55] *The Mercury* (Hobart), 2 February 1872, <http://trove.nla.gov.au/ndp/del/article/8921624>, accessed 6 April 2011.

[56] [Kinchin 2011], a *Sunday Times* article on academic crowdsourcing – featuring *Transcribe Bentham* – caused an increase in traffic to the Transcription Desk, and recruitment of several new regular transcribers. Sadly, the *Times* pay-wall prevented more widespread dissemination of the article.

[57] Three of the original seven “super-transcribers” continue to transcribe regularly, joined by three recruited after publication of [Kinchin 2011]. Another new super transcriber began taking part after finding that the *Dickens Journals Online* (<http://www.djo.org.uk/>) project was oversubscribed, and when the *DJO* team provided suggestions of other crowdsourcing projects in which to take part. Finally, an eighth regular participant began transcribing in mid-November 2011.

[58] See note 5.

[59] [Romeo and Blaser 2011] note that despite the risk that entering weather observations for *Old Weather* might “be considered ‘dry’” or repetitive after a short time”, volunteers latched onto and became engrossed in the human stories contained within the logs. Entering weather details from the logs as part of the *Old Weather* project, became, for some volunteers, a means to “follow the stories of vessels and people through to the end”. On 11 April 2011, an episode of the BBC popular science programme, *Bang Goes the Theory*, examined the rise of “Citizen Science” and paid particular attention to the *Zooniverse* projects.

Works Cited

- Achaleke et al 2011** Achaleke, B., G. Harwood, A. Koblin, L. Yan and T. Peixoto. “Guinea Pigs and Apples: Statement of the Digital Communities Jury”. In H. Leopoldseder C. Schöpf and G. Stocker, eds., *Prix Ars Electronica International Compendium: CyberArts 2011*. Ostfildern: Hatje Cantz, 2011. pp. 202-211.
- Beals 2010** Beals, M. “Review: Transcribe Bentham: a Participatory Initiative”. *Higher Education Academy* (2010). http://www2.warwick.ac.uk/fac/cross_fac/eahistory/resources/rv_beals_transcribebentham_20101026.pdf.
- Bennett 2010a** Bennett, T. “The Pursuit of Happiness: Jeremy Bentham, David Cameron and the Principle of Utility”, in *The Behaviour Guru Blog* (2010): <http://behaviourguru.blogspot.com/2010/11/pursuit-of-happiness-jeremy-bentham.html>.
- Brumfield 2011** Brumfield, B. *The Year of Crowdsourcing Transcription. The Collaborative Manuscript Transcription Blog*. 2011. <http://manuscripttranscription.blogspot.com/2011/02/2010-year-of-crowdsourcing.html>.
- Causer, Tonra and Wallace 2012** Causer, T., J. Tonra and V. Wallace. “Transcription Maximized; Expense Minimized? Crowdsourcing and Editing The Collected Works of Jeremy Bentham”. *Literary and Linguistic Computing* (2012).
- Chester Chronicle 2010** Chester Chronicle. “Queen’s School Students Involved in Transcribe Bentham Project”. *The Chester Chronicle* (2010). <http://www.chesterchronicle.co.uk/chester-news/local-chester-news/2010/10/07/queen-s-school-students-involved-in-transcribe-bentham-project-59067-27418197/>><http://www.chesterchronicle.co.uk/chester-news/local-chester-news/2010/10/07/queen-s-school-students-involved-in-transcribe-bentham-project-59067-27418197>.
- Cohen 2010** Cohen, Patricia. “Scholars Recruit Public for Project”. *New York Times* (December 27 2010). <http://www.nytimes.com/2010/12/28/books/28transcribe.html?pagewanted=all>.
- Cunnane 2010** Cunnane, S. “As a Discipline, We Suck Online”. *Times Higher Education* (July 15 2010). <http://www.timeshighereducation.co.uk/story.asp?sectioncode=26&storycode=412549&c=1>.
- Dunning 2011** Dunning, A. *Crowdsourcing and Variant Digital Editions: Some Troubles Ahead. The JISC Digitisation Programme Blog*. 2011. <http://digitisation.jiscinvolve.org/wp/2011/07/18/crowdsourcing-and-variant-digital-editions-some-troubles-ahead/>.
- Elken 2011** Elken, M. *Crowdsourcing – A New Way to do Scholarly Work? The Higher Education Development Association Blog*. 2011. <http://uv-net.uio.no/wpmu/hedda/2011/03/16/crowdsourcing-a-new-way-to-do-scholarly-work/>.

Family Search Indexing Family Search Indexing Project. *Family Search Indexing Project*.
<https://www.familysearch.org/volunteer/indexing>.

Haythornthwaite 2009 Haythornthwaite, C. "Crowds and Communities: Light and Heavyweight Models of Peer Production". In *Proceedings of the 42nd Hawaiian Conference on System Sciences*. Waikola: IEEE Computer Society, 2009. pp. 1-10.

Holley 2009 Holley, R. *Many Hands Make Light Work: Public Collaborative OCR Text Correction in Australian Historic Newspapers*. Canberra: National Library of Australia, 2009.

Holley 2010 Holley, R. "Crowdsourcing: How and Why Should Libraries Do It?". *D-Lib Magazine* 16: 3-4 (2010).
<http://www.dlib.org/dlib/march10/holley/03holley.html>.

Kinchen 2011 Kinchen, R. "One Stir, Then I'll Discover a Galaxy". *The Sunday Times* (September 11 2011).
<http://www.thesundaytimes.co.uk/sto/newsreview/features/article772703.ece>.

Moyle 2011 Moyle, M., J. Tonra and V. Wallace. "Manuscript Crowdsourcing: Transcribe Bentham". *LIBER Quarterly* 20: 3-4 (2011), pp. 347-356. http://liber.library.uu.nl/publish/issues/2010-3_4/index.html?000514.

Mroz 2010 Mroz, A. "Leader: Members-Only Code is Outdated". *Times Higher Education* (2010).
<http://www.timeshighereducation.co.uk/story.asp?sectioncode=26&storycode=412539&c=2>.

Old Weather Project Old Weather Project. *Old Weather Project*. <http://www.oldweather.org/>.

Oomen 2011 Oomen, J., and L. Aroyo. "Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges". Presented at *Fifth International Conference on Communities and Technologies, Brisbane*, sponsored by (June 29 - July 2 2011). <http://www.cs.vu.nl/~marieke/OomenAroyoCT2011.pdf>.

Organisciak 2010 Organisciak, P. *Why Bother? Examining the Motivations of Users in Large-Scale Crowd-Powered Online Initiatives*. Thesis, University of Alberta: 2010.

Phenology Survey 2010 North American Bird Phenology Programme. *North American Bird Phenology Programme User Survey (2010)*. 2010. http://www.pwrc.usgs.gov/bpp/SatisfactionSurveyReport2010/Satisfaction_Survey_2010.BK.html.

Raddick et al 2010 Raddick, M.J., A.S. Szalay, J. Vandenberg, G. Bracey, P.L. Gay, C. Lintott, P. Murray and K. Schawinski. "Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers". *Astronomy Education Review* 9: 1 (2010). <http://dx.doi.org/10.3847/AER2009036>.

Romeo and Blaser 2011 Romeo, F., and L. Blaser. *Bringing Citizen Scientists and Historians Together. Museums and the Web 2011: Proceedings*. Edited by J. Trant and D. Bearman. Toronto: 2011.
http://conference.archimuse.com/mw2011/papers/bringing_citizen_scientists_historians_together.

Schofield 2009a Schofield, P., *Bentham: A Guide for the Perplexed*. London: Continuum Books (2009).

Shaw 2010 Shaw, M. *Transcribe Bentham -- A Crowdsourcing Experiment*. *The British Library Growing Knowledge Blog*. 2010. <http://britishlibrary.typepad.co.uk/growingknowledge/2010/09/transcribe-bentham-a-crowdsourcing-experiment.html>.

War Department Papers War Department. *Papers of the War Department, 1784-1800 Project*.
<http://wardepartmentpapers.org/index.php>.

Winters 2011 Winters, J. *Digital Editions and Crowdsourcing*. *IHR ReScript Project Blog*. 2011.
<http://rescriptihr.blogspot.com/2011/06/digital-editions-and-crowdsourcing.html>.



This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.