EXTRA-ORGANIZATIONAL LEARNING: LEARNING BEYOND ORGANIZATIONAL BOUNDARIES

Montserrat Prats López[1] (m.pratslopez@vu.nl), Hans Berends[1] (j.j.berends@vu.nl), Marleen Huysman[1] (m.h.huysman@vu.nl), Maura Soekijad[1] (m.soekijad@vu.nl), Tim Causer[2] (t.causer@ucl.ac.uk), Melissa Terras[2] (m.terras@ucl.ac.uk), and Kris Grint[3] (kcg4@st-andrews.ac.uk).
1. Vrije Universiteit Amsterdam, The Netherlands; 2. University College London, UK; 3. University of St Andrews, UK.

Abstract

Organizational scholars have studied learning processes within and across organizational boundaries. In this paper, we introduce the concept of extra-organizational learning to explain the mechanisms that enable learning among external individuals who are not bound to the focal organization by an employment contract and yet their tasks benefit the focal organization. We study the crowdsourcing project *Transcribe Bentham*, where the crowd transcribes documents of historical importance, and show how participants learn by doing and by socially interacting with experts from the organization. We explain how this process of extra-organizational learning differs from known processes of inter-organizational learning and crowdsourcing as 'distant search' and we discuss implications for the organizational learning literature.

Key words: Organizational learning, extra-organizational learning, crowdsourcing, digital humanities, Transcribe Bentham.

Montserrat Prats López Msc
VU University Amsterdam
De Boelelaan 1105, 5th floor
1081 HV Amsterdam, The Netherlands
Telephone: +31 (0)20 5983660
Email: m.pratslopez@vu.nl

INTRODUCTION

Organizational learning is a process of change in organizational knowledge and behaviour that contributes to organizational performance (Argote & Miron-Spektor, 2011). According to the 4I framework, an organization learns as new individual insights are shared and transformed in interaction with others and ultimately embedded in organizational practices, routines, and systems (Crossan, Lane & White, 1999). Organizational learning does not always stop at the boundaries of an organization and may include external partners (Crossan, Maurer, & White, 2011; Easterby-Smith, Lyles, & Tsang, 2008).

A rich body of literature has been developed on organizational learning across boundaries (Easterby-Smith et al., 2008) mainly focusing on learning from or with other organizations. Recently, the interest in learning from external individuals has gained some attention, as organizations increasingly learn through distant search (Afuah & Tucci, 2012), by inviting individuals to solve problems. We argue that it may also be in an organization's own benefit to encourage learning by external actors on which the organization depends (Dyer & Nobeoka, 2000; Turner & Rindova, 2012).

In this paper, we study processes where learning extends to individuals beyond the boundaries of an organization, with the specific objective to benefit organizational performance. We refer to this process as *'extra-organizational learning'*. Key to the concept of extra-organizational learning is that learning takes place by individuals not employed by the organization. Extra-organizational learning does not focus on intra-organizational actors and processes but on external individuals who perform tasks that serve as input for the focal organization. It is, therefore, an extension of intra-organizational learning. In the management literature, we find several examples of extra-organizational learning for instance in service organizations, where the quality of the service partly depends on users of the service learning how to perform their role or use the service (Turner & Rindova, 2012).

So far, we know little about how extra-organizational learning actually occurs. Given the fact that external individuals do not have an employment contract, are less subject to organizational control, and cannot be easily socialized into the organization, common mechanisms for learning in organizational settings may not be applicable. Therefore, in this paper we set to answer the following question: what are the mechanisms that facilitate extra-organizational learning?

To uncover the mechanisms that facilitate extra-organizational learning, we carried out an empirical case study involving crowdsourcing applied to academic research. Extra-organizational learning is crucial in this context because participants need to learn how to make their contributions according to academic quality standards. We examined an influential crowdsourcing case in the field of humanities: the *Transcribe Bentham* project initiated by University College London (UCL). Through the mixed-methods study of this case, we have identified a condition for extra-organizational learning (retention), and two learning mechanisms (guidelines & feedback), that enable learning-by-doing and ensure

the quality of contributions. In the following pages we discuss intra- and inter-organizational learning and how extra-organizational learning relates to them. After reviewing the literature, we introduce our research setting and case, follow up with the findings of our study and conclude by discussing their implications for theory and practice.

THEORETICAL FRAMEWORK

Learning is a social and multi-level phenomenon defined as the *'interplay between social competence and personal experience'* (Wenger, 2000: 227). Individuals gain experience by seeking or being exposed to similar or different situations over time. New insights acquired through experience are interpreted in the context of everyday (working) life (Crossan et al., 1999; Zietsma, Winn, Branzei, & Vertinsky, 2002). As interpretations of new experiences or insights are shared with others, it becomes clear how close these new insights are to socially established norms and practices, and whether learning at the individual and group level actually takes place (Wenger, 2000). Sharing experiences with others often involves joint experimentation or exploration, and using its results to adapt interpretations (feedback) or to integrate new beliefs into group activities (Crossan et al., 1999; Zietsma et al., 2002). Learning becomes 'organizational learning' when new insights are embedded into the organization's systems, practices and routines, thus when new beliefs and behaviours are institutionalized (Crossan et al., 1999).

The 4I framework of organizational learning (Crossan et al., 1999) captures the multi-level dynamics of organizational learning stretching out over individual, group, and organizational levels. According to the 4I framework, learning involves intuiting new possibilities, interpreting these individual intuitions, integrating shared group understandings, and institutionalizing them into organizational norms and routines. Altogether, these processes constitute a feed forward movement of learning from individual to group learning, constituting the exploration of new possibilities. At the same time, institutionalized knowledge may feed-back to group and individual levels for exploitation. Exploration is in constant tension with the exploitation of institutionalized existing knowledge, for example by competing for organizational resources (Crossan et al., 1999; Berends & Lammers, 2010).

On the one hand, organizational learning is about exploring and gaining new and diverse knowledge. On the other hand, organizational learning is also about exploiting existing knowledge and becoming more competent and reliable (Holmqvist, 2003; March, 1991). It is generally accepted that as individuals, groups and organizations gain experience in the execution of a task, they are likely to become more effective in performing that task and deliver better results. This is the principle behind learning-by-doing or the learning curve, time and errors *"decrease at a decreasing rate as experience is gained with the task"* (Argote, 2013). Even when the accumulation of experience does not lead to improved outcomes (Huber, 1991), experience remains an essential element of learning at any level of analysis, or in the words of Argote and Miron-Spektor (2011: 1124) *"experience is what transpires in the organization as it performs its task"*.

As organizations specialize and become more complex, they increasingly collaborate with other organizations to innovate and achieve competitive advantage (Hislop, 2009). These collaborations are based on sharing single organizational experiences or creating joint experiences, thus learning from or with each other (Holmqvist, 2003). This means that learning does not stop at the boundaries of an organization (Crossan et al., 2011). A clear example of inter-organizational learning is that between a manufacturer and its suppliers. Dyer and Nobeoka (2000) studied the practices that Toyota developed to facilitate inter-organizational learning, which contribute to the high productivity achieved by the company and its network of suppliers.

Inter-organizational learning literature has contributed to the field by focusing on the factors that facilitate or hinder learning across organizational boundaries. The factors usually studied refer to the characteristics of the learners, the sources of knowledge, the relationships in the process of learning, and the type of activities that are learned (Easterby-Smith et al., 2008; Ingram, 2005). Researchers studying inter-organizational relationship dynamics have looked at different types of relationships (i.e. competitive, supplier, alliance) and how relational factors influence learning in these relationships (Argote, Denomme & Fuchs, 2011; Easterby-Smith et al., 2008; Ingram, 2005). Learning mechanisms differ depending on the type of inter-organizational relationship (Easterby-Smith et al., 2008) and include: routine based mechanisms, such as meetings, work groups, training or transferring employees; the organizational orientation to collaborate with external partners; inter-organizational shared norms; informal communication; and the actions of boundary-spanners (Easterby-Smith et al., 2008; Knoppen, Sáenz, & Johnston, 2011).

The study of learning beyond organizational boundaries has mainly focused on learning in inter-organizational relationships. Only recently has the interest in learning from individuals outside organizational boundaries increased, as the widespread access and low cost of new technologies allow organizations to reach and learn from a great diversity of individuals with distant but relevant knowledge and skills (Afuha & Tucci, 2012). This is mainly done to tap into the existing and diverse knowledge of the crowd to support learning through exploration. Just like in intra- and inter-organizational settings (Holmqvist, 2003), learning from individuals beyond organizational boundaries can be both explorative and exploitative. Organizations do not only search for new insights from the public, they also open-up their routines and involve individuals to contribute for the benefit of the organization. While learning from individuals outside formal boundaries has been studied as a phenomenon related to exploration (Afuah & Tucci, 2012), learning by outsiders for exploitation has only sporadically been analyzed (Dyer & Nobeoka, 2000). In this paper, we focus on knowledge exploitation and the learning of external individuals and groups.

In an intra-organizational context, mechanisms to support feedback learning from the organization level to groups and individuals rely on embedding new knowledge into organizational norms and routines (Crossan et al., 1999; Berends & Lammers, 2010). External individuals are not familiar with the organization's way of working, therefore, incorporating new knowledge to existing organizational norms and routines is not likely to stimulate their learning. In an inter-organizational setting, learning mechanisms are

meant to support learning from business partners, which are not the external individuals we are focusing on here. Hence, learning mechanisms, such as relying on shared norms, creating workgroups or transferring employees (Easterby-Smith et al., 2008; Knoppen et al., 2011), might not be appropriate to support external individuals' learning, because they are not bound to the organization by an employment contract or business relationship agreement to control their activities.

We define *extra-organizational learning* as the process by which individuals, not bound to the organization with a business or employment contract, become competent in the performance of a task, the outcome of which benefits and influences organizational routines, thereby contributing to the organizational objectives. The quality of the task is essential to achieve organizational goals, whether these are better consumer experiences (Frei, 2006) or high quality data to support scientific research (Wiggins & Crowston, 2011). Figure 1 shows extra-organizational learning as a process that takes place beyond organizational boundaries involving external individuals and groups. The process is presented as mirroring the feedback learning process in the 4I framework on organizational learning (Crossan et al., 1999; Zietsma et al., 2002).

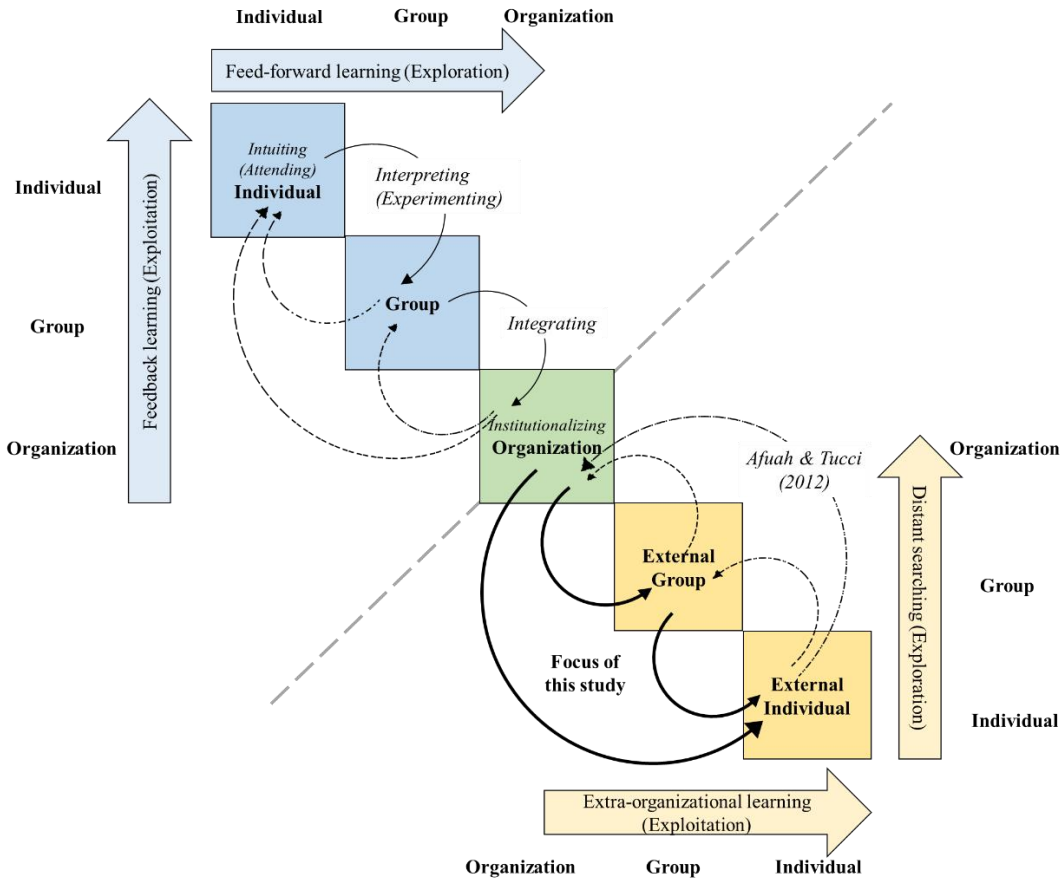

Figure 1. Extra-organizational learning mirroring the 4I framework (Crossan et al., 1999; Zietsma et al., 2002)

We find some examples of extra-organizational learning in the management literature, especially in studies about (often digitally enabled) services. For instance, services where users perform part of the service work, such as: online booking of flight tickets; or self-service stores, where consumers choose products and carry them to the counter. The benefits for the organization are lower labour costs and the likelihood of a better experience or service quality for the customer (Frei, 2006, 2008). The latter, however, is partly determined by how well customers have performed their share of the task. This depends on the needs, preferences and effort put by customers on that task. But more importantly, it also depends on the capabilities required to execute that part of the service and the variability in customer skills.

While having skill variability is essential for exploration, reducing skill variability is at the core of exploitation. Skill variability can be reduced by targeting skilled people, training people, or accommodating the task to their skills (Frei, 2006). For example, Turner and Rindova (2012) show how waste management organizations train citizens, through garbage preparation rules, and simplify their tasks through consistent garbage collection times. We can say that reducing variability involves learning and requires specific mechanisms to support learning by external individuals for better organizational outcomes. Services represent one setting where extra-organizational learning takes place, but we argue that there are other situations where it applies, for instance crowdsourcing.

Crowdsourcing is a relatively new form of organizing work via the Internet that has been defined as: *"the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call"* (Howe, 2006). Despite the advantages of efficiency and knowledge diversity usually associated with crowdsourcing (Brabham, 2013), this form of organizing is not without challenges. This is particularly the case when tasks are knowledge intensive, such as in research contexts in private or public organizations (Franzoni & Sauermann, 2014). Individuals participating in a crowdsourcing project are not bound to the organization with an employment contract, that is, they can decide when, where and how they perform that particular task. The openness of crowdsourcing also means that everyone can decide whether to participate or not. This results in uncertainty for the organization about the number of participants, the rate of work, and the knowledge that these participants will bring to the project. Consequently, participants in crowdsourcing initiatives have considerable autonomy and power in their relationship with the organization.

Crowdsourcing often calls for extra-organizational learning. Organizations initiating crowdsourcing projects outsource tasks, previously carried out within organizational boundaries, to individuals who volunteer their time and knowledge to perform these tasks (Howe, 2006; Wiggins & Crowston, 2011). To ensure that contributions meet quality standards, the focal organization communicates task requirements and explains how participants should perform the task properly. By sharing this knowledge with them, opportunities for learning arise. That is, just like in service settings the organization trained its users, in a crowdsourcing context the organization educates participants. And, as these participants contribute by repeatedly performing the outsourced task they learn along the way, they learn by doing (Argote, 2013).

We know little about how the process of extra-organizational learning actually happens. Therefore, in the following pages we study how extra-organizational learning takes place and what mechanisms are used to support learning in the crowdsourcing project *Transcribe Bentham*.

METHODS

We intend to contribute to theory development and expand an existing organizational learning model (4I framework by Crossan et al., 1999) by studying the external organizational context, and in particular the new phenomenon of crowdsourcing. A case study approach allows us to examine this phenomenon as it happens and contribute to theory building (Gibbert & Ruigrok, 2010). Moreover, a mix of qualitative and quantitative data can give us a more accurate picture of it, than if we were to rely only on one type of data or the other (Edmondson & McManus, 2007). Therefore, in this study we adopt a mixed methods case study approach.

Research setting

We study extra-organizational learning in the context of scholarly crowdsourcing. Scholarly crowdsourcing projects are a suitable research setting because participants need to learn how to perform a task according to professional standards for their contributions to be valuable to researchers and to become input to scholarly practices. The field of humanities (e.g. history, literature) is particularly interesting because external individuals contribute not only in the data collection phase of research, but mainly in the knowledge intensive phases of data processing and analysis.

Our case study is the award-winning[1] project *Transcribe Bentham* initiated by University College London (UCL). The aim of this project is to transcribe the collection of original manuscripts written and composed by Jeremy Bentham, a famous British philosopher and jurist who lived in the late 18th and early 19th centuries (Causer, Tonra & Wallace, 2012).

The transcription of Bentham's work, with the intention of producing the authoritative edition of the *Collected Works of Jeremy Bentham*, started 60 years ago. Given the size of Bentham's manuscript corpus (about 75.000 manuscript pages), the limited funding and existing technology, it was not until 2010 that, with the help of the Internet, the online collaborative digital transcription of his writings began.

[1] *Transcribe Bentham* won an Award of Distinction from Prix Ars Electronica 2011 in the category 'Digital Communities'.

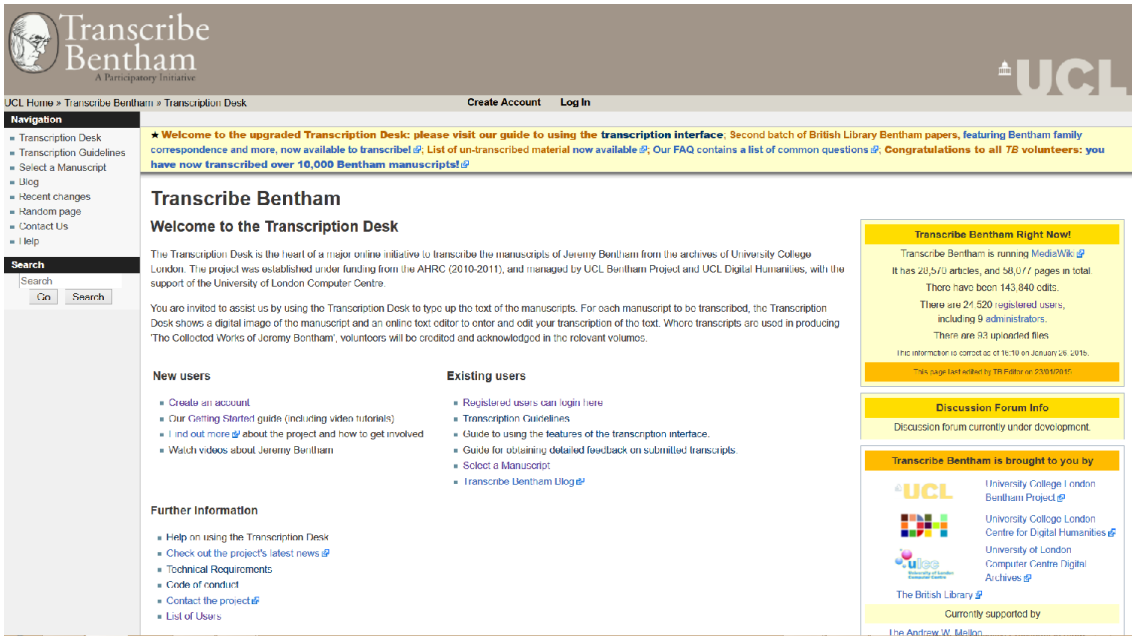

Figure 2. Home page of the Transcribe Bentham project

Project *Transcribe Bentham* aims at increasing the pace of producing transcriptions, which are used to contribute to the production of Bentham's Collected Works, and to be made available in UCL's digital Bentham Papers repository, to facilitate public access and searching.

The *Transcribe Bentham* team of scholars created an online environment, based on MediaWiki (see Figure 2). It allows people from all over the world to transcribe Bentham's work as voluntary participants, who do not receive any monetary rewards. The task of these external participants is to literally transcribe Bentham's manuscripts. This is done in the designated online environment, following certain standard rules, and encoding the text into TEI-compliant[2] XML. Figure 3 shows the MediaWiki transcription environment, with a scan of the manuscript on the right and the transcription done by one of the participants on the left. The project is managed by two highly dedicated UCL employees, who are responsible for the communication with participants, answering questions and stimulating participation. They also act as editors by reviewing and improving participants' contributions to ensure quality outcomes. To stimulate participation and friendly competition, they have also included a ranking of top contributors on the project site.

[2] TEI stands for *Text Encoding Initiative* and refers to the community of scholars that have developed standards to program texts in digital form. Source: http://www.tei-c.org/index.xml

Transcribe Bentham is an ideal case to study how extra-organizational learning takes place and thus helps us to answer our research question.

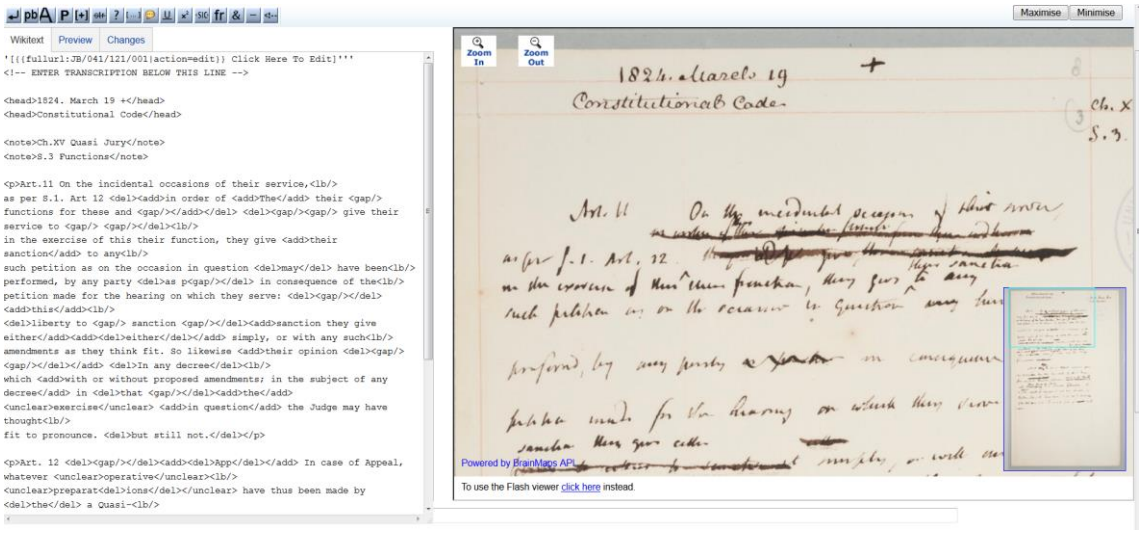

Figure 3. Example of a complex manuscript (on the right) and its transcription and encoding (on the left)

Data collection and analysis

To understand how extra-organizational learning takes place, we gathered quantitative and qualitative data.

The collection of quantitative data was carried out by the co-authors employed in *Transcribe Bentham*. The project started in 2010, but the data for this study was collected in the period from October 1st, 2012 to June 27th, 2014. It includes data from all the participants in that period (78 people) who together submitted a total of 4.303 pages in that period, which were checked and approved by *Transcribe Bentham* staff. For each transcribed page, the collected data includes: number of words; level of difficulty (readability) of the handwriting; number of alterations (changes) that the editors made to participants' transcriptions; and a code to identify individual transcribers (anonymized). The level of difficulty or readability of the handwriting in each page has been classified as easy, moderate and difficult by the UCL experts. The classification is based on the period that Bentham wrote, that is, as he grew older his handwriting worsened, thus manuscripts written between 1770 and 1789 are easier to read than those written in 1832.

These quantitative data allowed us to examine whether learning took place, and whether it benefitted the organization in terms of quality contributions. We first looked at the level of participation and activity over the studied period of time. For that purpose, we used a Lorenz curve as suggested by Sauermann & Franzoni (2015) to show the distribution of contributions among participants. For those who transcribed more than 100 pages (whom we call "super-transcribers"), we investigated how the number of alterations changed overtime, as an indication of performance in terms of quality. Moreover, these data also showed the accumulation of experience of these super-transcribers. The accumulation of

experience is known as learning-by-doing (Adler & Clark, 1991) and learning-by-doing is often illustrated with experience curves. An experience curve shows the improvement of outcomes as a task is repeated overtime, that is, as experience is accumulated (Argote, 2013). Then, we further looked into the data to explain differences between super-transcribers and other participants. We examined the quality of the first attempts at transcribing and compared the means among different types of participants and different levels of manuscript difficulty.

Assessing learning based on performance improvements can be challenging, because other factors besides learning can affect performance (Argote & Miron-Spektor, 2011; Easterby-Smith et al., 2008). Therefore, we used qualitative data to corroborate and clarify our quantitative findings and to give insight in the underlying mechanisms of the learning process. Qualitative data helped us to understand what participants consider important activities for learning. The qualitative data consist of an archive of 4,489 e-mail notifications, notes from meetings and informal communication with the project staff, and interviews with 4 super-transcribers. The interviews were completed in September 2014 by the first author. During the interviews, participants were asked to explain how they perform the task, what makes a good transcription, and what would they recommend us to do if we were to contribute for the first time. Notes and interview transcriptions were stored and analysed in the qualitative research software tool ATLAS.ti.

FINDINGS

In the following pages, we first explain what quality means in this context and why it is important. Then we show evidence of extra-organizational learning and how learning results in better quality. Finally, we discuss the mechanisms that the organization (UCL) uses to exploit institutionalized knowledge and to facilitate extra-organizational learning.

Quality

Quality, as defined by the editors of *Transcribe Bentham*, includes both the textual accuracy and the encoding consistency of the transcriptions. *Accuracy* is related to the literal transcription of the handwritten text. Transcriptions need to be a truthful representation of the original text if they are to be valuable input for historical and textual research, and for public searching. In *Transcribe Bentham*, transcriptions must even include Bentham's mistakes, deletions and unusual or idiosyncratic spellings. Encoding *consistency* refers to the correct usage of the TEI-compliant XML codes. The uniformity in coding is important for a proper representation of the structure and features of the text, namely: headings, paragraphs, side notes, deletions, etc.

Interviews with super-transcribers show that participants understand the importance of quality as explained above. They understand that transcriptions need to be accurate:

> *"It's just trying to replicate and give the fullest possible representation of the original author's words and intentions as they have been placed on the parchment really, it's just ... just to give it a fair representation of what they wanted to put into words, what they put into the manuscript themselves."*

Transcribers also realize that transcriptions need to be consistent and complete:

> *"A lot of times it's just a matter of completion. How much of the text have you transcribed, did you use the html correctly, […]be able to look at the manuscript and say: 'yep, I used the headings button right that time' or 'oh! I should fix that, I put that on the wrong block of text' kind of thing. So, it's just kind of a matter of completion, giving things as close as transcribed as you possible can…".*

E-mail messages received with submitted transcriptions demonstrate how seriously participants take their tasks. They often apologize when they have not been able to complete a transcription or when they feel that the manuscript was difficult, telling *Transcribe Bentham* staff that there are *"Still lots of gaps and guesses. Sorry!"*, or that *"I've done the best I can with this one…"*. The conscientiousness of transcribers when going about the task is clear, since they sometimes feel it necessary to *"apologize in advance for any mistakes or oversights"*.

Quality is important because the transcriptions are used as input for scholarly research. When participants submit higher quality transcriptions, *Transcribe Bentham* editors need to invest less time reviewing and correcting the submitted transcriptions. We measure quality in terms of the number of alterations (or changes) that the editors need to make to a transcription. Figure 4 shows that the time spent checking transcriptions is greater when a higher percentage of alterations per word is required to ensure that the transcription is according to standards and is accepted by the project staff.

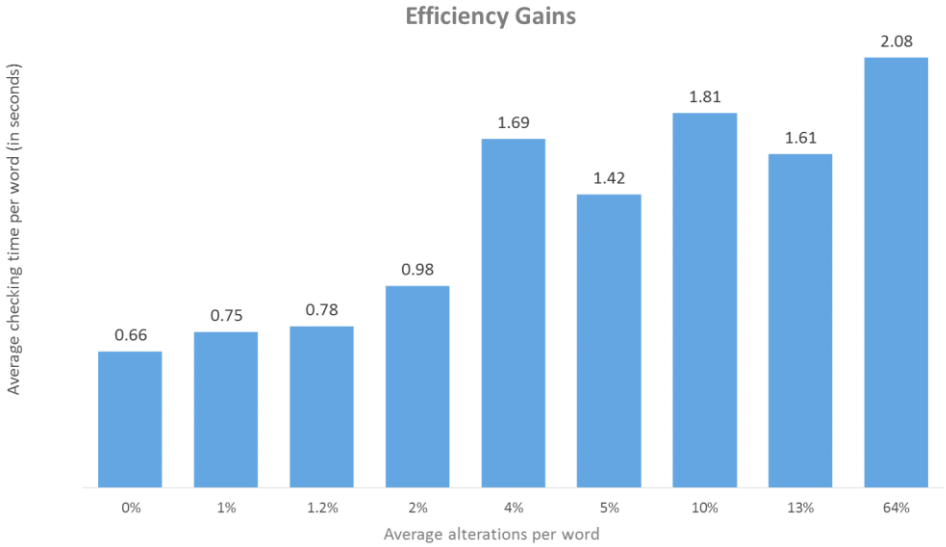

Figure 4. Efficiency gains from correcting transcriptions with better quality

The organization benefits from efficient reviewing of transcriptions when the quality increases, that is, when *Transcribe Bentham* staff need to make few alterations to the transcriptions before accepting them. For example, editors can review and correct an average of 18 transcriptions in one hour if they require fewer than 50 alterations each, while they can review and correct more or less 3 in one hour if more than 50 alterations need to be made to each transcription.

Evidence of extra-organizational learning

At the *individual level*, learning takes place as people repeatedly perform a task. Therefore, we first examine how many people participate, that is, perform transcriptions in *Transcribe Bentham*. The quantitative data shows that 78 people contributed at least once in the 21 months period from 01/10/2012 to 27/06/2014. From all participants, 11 people contributed 91% of the transcriptions. This is illustrated in the Lorenz curve below. The distribution of contributions follows the Pareto principle in the sense that a few people contribute the most. This is not only common in crowdsourcing initiatives (Owens, 2013) but also in online communities in general (Faraj, Jarvenpaa & Majchrzak, 2011).

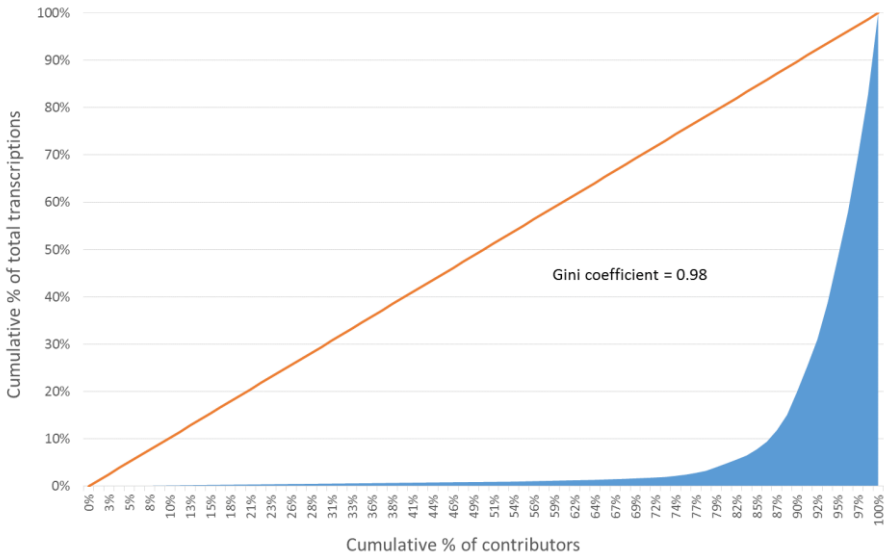

Figure 5. Lorenz curve indicating the unequal distribution of participation

We refer to the 11 participants who contributed most to the project as super-transcribers. They participated in the project for a longer period of time than other participants and they submitted more transcriptions. They also were more active in terms of the number of contributions done per active day of transcribing (2.7 on average) and had shorter breaks or periods of inactivity (2.3 days on average) compared to other participants (2.2 transcriptions per day and 15.7 days in between activity on average).

Out of the 78 people in total who contributed to *Transcribe Bentham* from 1-10-2012 to 27-06-2014, 63 started participating during that period, while the other 15 had joined before October 2012. In our analysis, we focus on the 63 people who started in this period,

because we have complete data about their activity. Table 1 includes the measures of participation and activity for these people. In the observation period, 38 people contributed one transcription; 19 people transcribed more than one page but less than 100; and 6 super-transcribers transcribed 100 pages or more. Altogether, these participants contributed 1,618 transcriptions in 508 days with an average of 3.2 transcriptions per active day.

Participation and activity Measures		Total	Transcribers 1 page	Transcribers 2 to 100 pages	Super-transcribers >100 pages
Nr. of people	Total	78	42	25	11
	Started in research period*	63	38	19	6
	First transcribed page	63	38	19	6
Nr. of transcriptions	Total	4,303	42	363	3,898
	Started in research period*	1,618	38	233	1,347
	First transcribed page	63	38	19	6
Share of transcriptions	Total	100%	1%	8%	91%
	Started in research period*	100%	2%	14%	83%
	First transcribed page	100%	60%	30%	10%
Nr. of active days contributed	Total	1,657	42	166	1,449
	Started in research period*	508	38	84	386
	First transcribed page	1	1	1	1
Average transcriptions per active day	Total	3	1	2	3
	Started in research period *	3	1	3	3
	First transcription	1	1	1	1
Average duration of breaks (in days) **	Total	10.8	-	15.7	2.3
	Started in research period*	8.6	-	11.5	2.5
	First transcribed page	-	-	-	-

Table 1. Measures of participation and activity in Transcribe Bentham
** Participants who started transcribing in the period from 1-10-2012 to 27-06-2014.*
*** Based on people who contributed more than one transcription in different days.*

Next to the level of participation and activity, we also looked at participants' performance in terms of the number of alterations that were made to their contributions by *Transcribe Bentham* staff (see Table 2). A first look at the data shows that the transcriptions submitted by super-transcribers needed on average fewer alterations than those submitted

by other participants. A Kruskal-Wallis[3] test showed the differences in the average number of alterations between these three groups were statistically significant. That points at the existence of a relationship between the quality of the transcriptions in terms of alterations, and the level of participation in the project. This suggests the existence of a learning-by-doing process among super-transcribers, based upon all the available data. However, an explanation for why super-transcribers perform better than the other transcribers, could be their prior knowledge: maybe they were already more skilled than the other transcribers when they started. They might have had more experience or better skills in deciphering handwritten texts, which might have stimulated them to do more transcriptions, whereas less skilled participants may have experienced more difficulties and might be more inclined to quit after one transcription. Therefore, we looked into the details of the first transcriptions submitted by those participants who joined the project during the period of study. We compared the first transcriptions of those who became super-transcribers and those who did not (see Table 2). The average number of alterations in the first transcription between these groups is not statistically different. The quality or performance of the first transcription does not seem to be related to becoming a super-transcriber. Therefore, we rule out prior knowledge as an alternative explanation for better performance by super-transcribers.

Learning Measures		Total	Transcribers 1 page	Transcribers 2 to 100 pages	Super-transcribers >100 pages
Nr. of alterations	Total	12,941	398	2,423	10,120
	Started in research period*	6,729	386	1,644	4,699
	First transcribed page	785	386	309	90
Average alterations per transcription	Total	3.0	9.5	6.7	2.6
	Started in research period*	4.2	10.2	7.1	3.5
	First transcribed page	12.5	10.2	16.3	15.0
Average alterations per word	Total	0.04	0.05	0.03	0.01
	Started in research period*	0.05	0.07	0.04	0.01
	First transcribed page	0.06	0.05	0.07	0.05

Table 2. Learning measures in Transcribe Bentham

** Participants who started transcribing in the period from 1-10-2012 to 27-06-2014.*

Another potential alternative explanation for super-transcribers performing better than the other groups of participants could be the level of difficulty of the manuscripts they chose to transcribe. In general, the average number of alterations made to a transcription

[3] We compared the means in the number of textual alterations per word. The data is not normally distributed and the samples are small, therefore we performed a Kruskal-Wallis test.

will differ between manuscripts considered easy, moderate and difficult and we might expect the same in the first transcription.

Difficulty Level		Total	Transcribers 1 page	Transcribers 2 to 100 pages	Super-transcribers >100 pages
Easy	Total	48%	60%	56%	47%
	Started in research period*	51%	55%	58%	50%
	First transcribed page	51%	55%	42%	50%
Moderate	Total	34%	26%	36%	34%
	Started in research period*	32%	29%	35%	31%
	First transcribed page	33%	29%	53%	0%
Difficult	Total	18%	14%	7%	19%
	Started in research period*	17%	16%	8%	19%
	First transcribed page	16%	16%	5%	50%

Table 3. Selected difficulty levels
** Participants who started transcribing in the period from 1-10-2012 to 27-06-2014.*

However, the statistical analysis indicates no significant differences in the number of alterations made in the first transcription, whether this was easy, moderate or difficult. No significant differences were found either for each of the group of transcribers and the difficulty of the first transcribed page (Table 4).

Level of Difficulty First Transcription	**Mean**			**Sig.**
	Transcribers 1 page	**Transcribers 2 to 100 pages**	**Super-transcribers >100 pages**	
Easy	0.046	0.074	0.059	0.844
Moderate	0.057	0.071	-	0.833
Difficult	0.039	-	0.044	0.263

Table 4. Kruskal-Wallis test
We compared the means in the number of textual alterations per word for the three groups of transcribers and levels of difficulty.

A final possible explanation for super-transcribers performing overall better than other participants is that super-transcribers might have chosen easier manuscripts to transcribe over time. Figure 6 shows the manuscript difficulty chosen by super-transcribers in their first 100 pages. Up to their first 70th page, they select mainly easy and moderate manuscripts on which to work. After the 70th page, the percentage of easy manuscripts increases. Though after the 90th page the percentage returns to the prior level. In total, super-transcribers chose easy pages 50% of the time, moderate pages 31% of the time

and difficult pages about 19% of the time. Therefore, the difficulty of the manuscripts does not seem to explain their learning process.

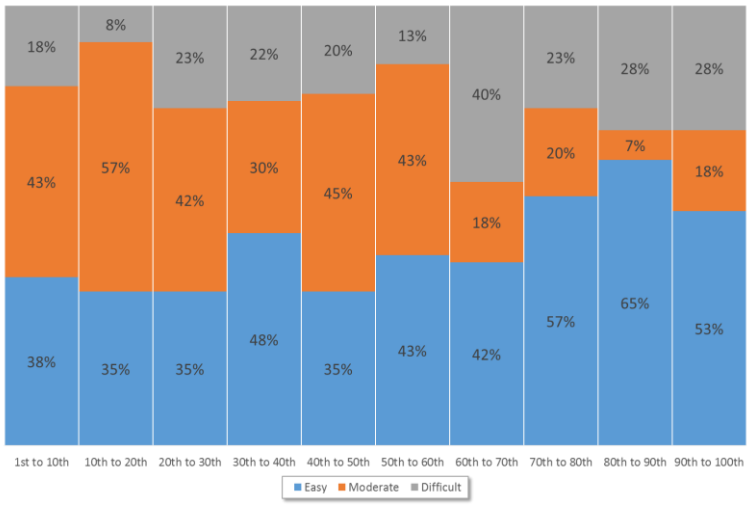

Figure 6. Percentage of easy, moderate and difficult manuscripts chosen overtime by 6 super-transcribers.

The overall performance of *Transcribe Bentham* participants is illustrated in figure 7. The figure shows the aggregated data for all participants in order of transcription. We see how overtime the average number of alterations per word declines as participants gain experience and transcribe new manuscript pages. The average percentage of alterations per word declined from about 4% in the first transcription to 1% after the 50th transcribed page.

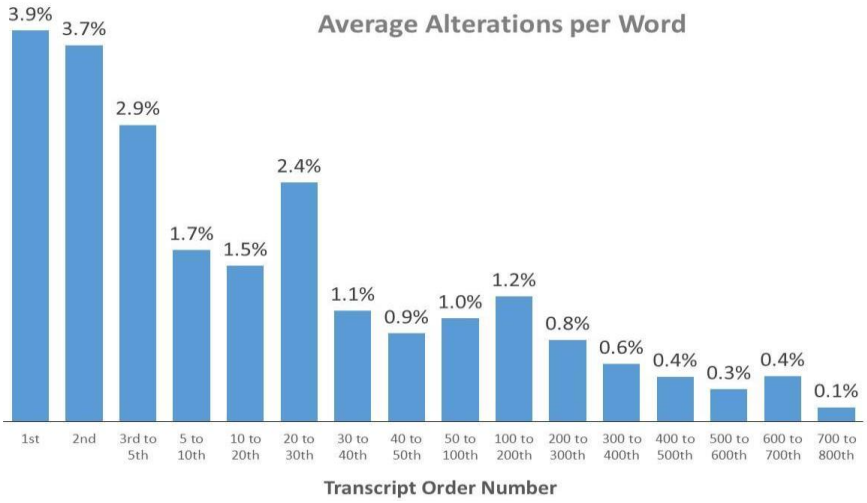

Figure 7. Average number of alterations as participants gain experience

This shows the importance of experience or learning-by-doing at the core of extra-organizational learning. The number of alterations made to the transcriptions declines as participants gain experience. This indicates that at the *individual level*, crowdsourcing participants outside organizational boundaries that perform a task for an extended period of time become better at it, they learn and deliver higher quality.

Mechanisms for extra-organizational learning

In this section, we examine the mechanisms used by the organization (UCL) to support extra-organizational learning.

Transcribing historic manuscripts has long been a key step in the research process of humanities scholars. Transcribing includes a number of *institutionalized* practices and rules that are widely used in the field. A 'diplomatic' transcription (i.e. literally copying handwritten text, and attempting to represent the features of the text) is one of the methods most commonly used. Though encoding text in TEI-compliant XML language is relatively new in the field of humanities (1994), it is now well established and can also be considered an institutionalized practice in the digitalization of textual data. Both institutionalized practices need to be exploited or shared with the crowd if they are to contribute to scholarly research. This process of exploitative learning takes place through several mechanisms.

1. Guidelines

As a crowdsourcing initiative, participation in *Transcribe Bentham* is fully Internet-based, using a customized MediaWiki environment known as the 'Transcription Desk', where scans of handwritten manuscripts are visible and transcriptions can be typed in. The Wiki also includes textual and visual guidelines for participation. These guidelines inform and instruct participants about the 'diplomatic' transcription method, point to important features of the manuscripts and explain how to use the encoding toolbar. The guidelines are simple and short, avoiding the use of jargon, and supported with a section on 'Palaeography Skills', a list of words commonly used by Bentham and some visual examples of his handwriting.

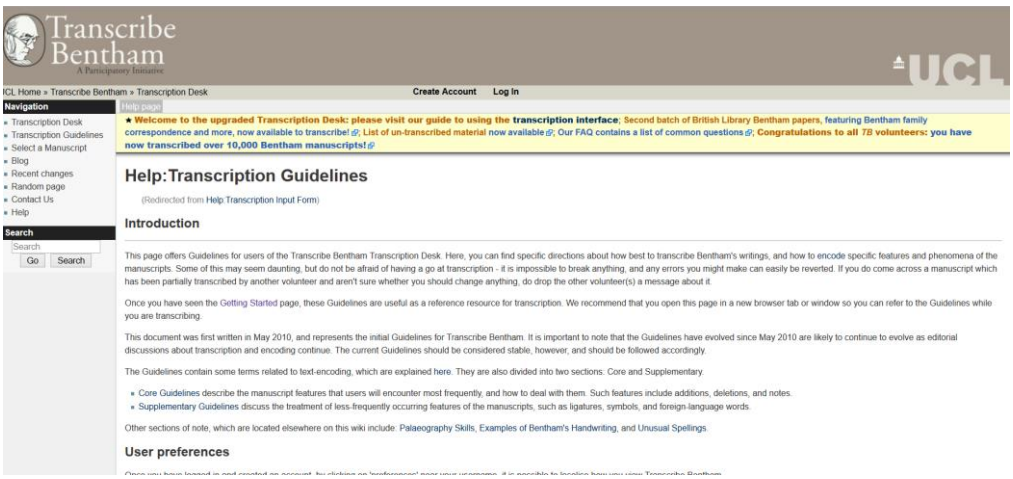

Figure 8. Transcription guidelines of Transcribe Bentham (retrieved on February 2015)

Participants are well aware of the availability of the guidelines. When asked what they did after registering for the first time, the guidelines are one of the first things mentioned:

> *"Basically just... registering is quite quick, you set up a short profile, there is some guidance on the method of transcription, it has to be put in a particular format. […] It gives*

you quite a thorough little guide of how to ... how it needs to be organized into digital form so that it can be useful to the project. And then you just kind of... you get through into it really…"

And when asked what a new participant should do to learn*:*

"Definitely read the transcription guidelines. They are right on the side bar and they are so helpful, I relied on them so much when I was just starting out, it will tell you exactly how to use the interface, like I mentioned earlier, there are also examples of [author]'s handwriting."

Though, not everyone is as keen to consult the guidelines, be they for *Transcribe Bentham* or any other project: *"… I think there was an instruction manual somewhere but I don't ever really use those, I just like to do it myself."*

2. Feedback

Another way by which the organization feeds institutionalized knowledge to the *individual level* is by giving participants feedback about their performance. The project staff members check the quality of the transcriptions by assessing accuracy and completeness. They revise the transcriptions to make sure there are as few as possible gaps or words marked as unreadable. After reviewing and correcting the transcriptions, they provide feedback to each participant. According to the project staff, feedback is important for the general participant experience, to keep the pace of transcription and maintain the quality of contributions.

Feedback is provided in two ways: through e-mail messages, the length of which depend on the number and type of alterations that were made; and through the revision history, which is a section of the Wiki that allows users to compare their submission with the project staff revised version. The e-mail messages are sent by project staff to inform participants about the specifics of alterations made. But when submitting a transcript, the transcribers can also include a short message in which they can ask questions, indicate issues with the manuscript, or request extra support from the project staff. The messages are submitted along with the transcript through the 'send a message to the editors' option embedded in the transcription area. The project staff members receive a message for each submitted transcription. They usually answer questions in the transcriber's board for sake of easiness. Transcribers' e-mails include comments to inform that it's their first transcription: *"That was my first attempt to transcribe Bentham. I'm not sure how successful. I would be grateful for any feedback".* They also point out that a page is blank, that there are words they could not transcribe: *"The first word in the heading eludes me"*, or express their difficulties with the text: *"There is a lot of guesswork in this transcription as the handwriting is rather difficult."* Sometimes they give up on some words and share their frustration with the editors: *"I'm afraid the handwriting has defeated me on this one."* But they also ask questions about a word: *"…the word I couldn't decipher on* [manuscript] *looks like* [word]. *Is that right?"* and about how to encode: *"… in a case where there is so much hard to see, should I continue to code the empty space and unknown, or leave it the tangle that I have made?"*.

Participants are aware that they receive feedback in the form of e-mail:

> *"... whenever you get a message back, I get them in my e-mail... they say 'thank you for your contribution, and sometimes they mention whether they needed to fix stuff. Cause usually when they needed to fix something they mention it. And otherwise it's just 'thank you for your contribution' and then you can go back to the page that you transcribed and you can see whether and what they changed."*

They are also aware that feedback on their work can be gained via the Transcription Desk's extensive revision history:

> *"I can see right under the production log that TB editor made a 57 character change to certain document, so if I hit to the difference button right next to it, I can see a page that says: 'differences between revisions of certain manuscripts' and I can see that where I added a gap of eligible text, they were able to actually see what it said and they transcribed it, added it in there. And it's just a very helpful tool to use"*

They acknowledge that the feedback received, either by e-mail or through the revision history, help them to become better at transcribing. Some participants are really keen to know what they did wrong: *"I'm quite anxious to see if I did anything wrong, quite keen not to give people any more work than they need to. […] It is very useful to know what you made wrong."* And they do acknowledge the positive effect of feedback on their learning: *"Because the first time that I did it, I didn't know what to do with that (click), cause you know you're still trying to learn, and I sent it to the editors and they sent me their revised version back, so that I could see what I needed to do, like which codes I needed to use where, what I should have done in a different way."*

3. Community building

With regard to the organization supporting community building or participants learning at a *group level*, project staff originally included a discussion forum in the Wiki. The forum was meant for giving feedback and to provide participants a virtual space to ask each other questions and share ideas. However, the discussion forum was barely used as a group tool and there was limited interaction among participants (Causer & Wallace, 2012). The topics discussed ranged from questions about the points system (i.e. to encourage participation people can get point for the tasks performed and be included in the leader board) to questions about transcribing certain words. In general, few other participants jumped in to offer answers or suggestions to questions; they rather waited for the project staff to respond. Moreover, technical issues resulted in the forum being unavailable since July 2013, though no complaints from the participants have been received about its unavailability.

Another way by which the Transcription Desk supports learning through community building is the social profile feature available to each registered user. The social profile allows participants to let others know about themselves or to share any favourite Bentham-related information and material with other transcribers. The social profile also includes a "message board", where participants can post public or private messages

to other project participants and add them as friends. The project editors post their standard acknowledgement messages publicly in each participant's board.

According to the project staff, there has been little communication among participants through the boards of which they are aware. Given that participants can send private messages, it could be possible that some interaction took place without the project staff knowing about it. However, based on the interviews held with super-transcribers, we conclude that the communication between participants is limited: *"Yeah, they have the messages of individual users and things like that, not really done so. You can sort of add them as a friend and things like that. […] No, never used any of those."* and, when it takes place, it mainly runs through other communication platforms like Twitter and Facebook: *"Okay, No, I haven't used that, no. Basically when I started doing it, I ... you don't want anything to distract you. Yeah. I have a friend that's on there but we just talk on Facebook to one another*."

Finally, *Transcribe Bentham* organized two events or information sessions in 2011, to further engage participants. However, this attempt to get together face-to-face was unsuccessful, as only one transcriber attended (Causer & Wallace, 2012). Such low attendance, in a project where participants are spread around the world, is perhaps unsurprising. Nevertheless, group level mechanisms to support learning through community building did not appear to be effective in this particular case.

To summarize, participants *'definitely read the transcription guidelines'*, but they also *'have a look at what other people have completed'*, *'try to find documents that are easiest to transcribe'*, thus *'finding quite a reasonable simple one and give it a go'*, and learn-by-doing, *'as you get more comfortable just work your way up to newer documents and those documents that are harder to understand'*. That is, the organization creates opportunities for learning through guidelines and feedback, and participants make use of them to learn.

DISCUSSION AND CONCLUSION

This paper contributes to the organizational learning literature in three ways: it expands the 4I framework by including extra-organizational learning; it uncovers and examines the specific mechanisms that facilitate extra-organizational learning; and shows how different theoretical traditions are merged in the extra-organizational learning process. In the following pages we discuss these contributions in more detail.

First, we contribute to the organizational learning literature by expanding the 4I framework (Crossan et al., 1999) to include extra-organizational learning, and by showing other ways through which organizational knowledge flows out of the organization into the external context (Argote & Miron-Spektor, 2011). This addition to the model shows that extra-organizational learning differs from inter-organizational learning, which has been modelled as an extra level in itself (Jones & Macpherson, 2006). Our paper considers all levels of analysis: organization, group and individual.

The management literature has studied multi-level learning within organizations (Crossan et al., 2011), but how learning crosses multiple levels in relationships between and beyond organizational boundaries has been largely overlooked. Organizational learning scholars have mainly focused on learning within or between organizations and less attention has been paid to the impact of learning beyond organizational boundaries. That is, learning by individuals in their roles of consumers, users or crowdsourcing participants, that directly affects the performance of organizations they interact with. By expanding the 4I framework to include individuals and groups beyond the boundaries of the organization, we theoretically and empirically contribute to the tree of knowledge about organizational learning (Crossan et al., 2011). We add the concept of extra-organizational learning, describe the instances in which it takes place, and explain the mechanisms that facilitate its occurrence.

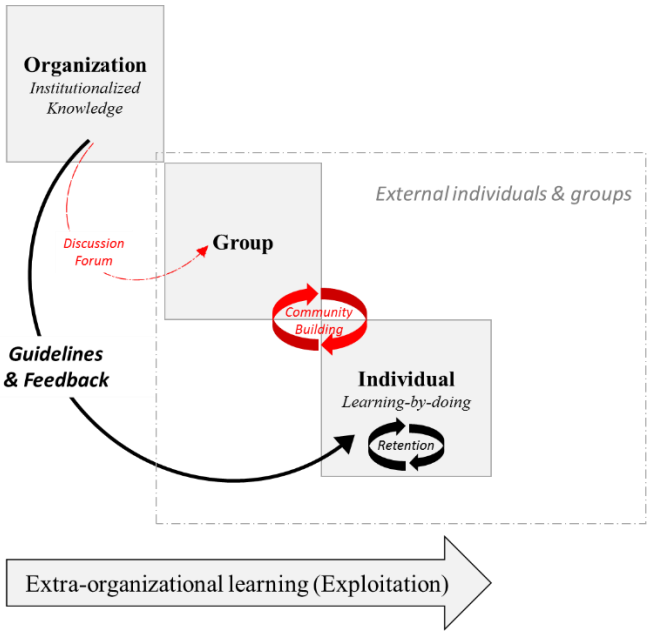

Figure 9. Mechanisms supporting extra-organizational learning

Second, we examine in detail the mechanisms that facilitate extra-organizational learning (see Figure 9). Our findings show that exploiting institutionalized knowledge enables external individual learning. The organization embeds its rules and norms into guidelines and feedback messages. Guidelines are the means by which institutionalized knowledge is codified and shared with individuals outside the organization's boundaries. Project staff provide feedback to participants about their performance based on institutionalized norms. *Guidelines* and *feedback* are, therefore, the mechanisms that facilitate extra-organizational learning at the *individual level*.

We have shown that participants need to perform a task for a certain period of time in order to learn from experience. Consistent with the literature on online communities (Faraj et al., 2011), participation varies over time and generally, only a smaller number of people become core contributors. Therefore, we conclude that one condition needed for extra-organization learning is ensuring higher participant (or user) *retention*, that is, keeping individuals engaged with the task for longer periods of time.

Mechanisms supporting learning from the organizational to the *group level* have proven to be a challenge in the studied case. Despite providing a discussion forum and other social interaction features, *Transcribe Bentham* remains a crowdsourcing initiative with a small community of unconnected super-transcribers, supported by project staff. The literature about online communities indicates that the levels of participation and retention might differ depending on the purpose of a community (Ren et al., 2012: 846, 858). Given the purpose of *Transcribe Bentham*, participants are unlikely to have joined for its social aspects, but for the content and the challenge and intrinsic enjoyment of the task at hand. Hence, group level learning is not likely to occur in crowdsourcing initiatives involving independent individual tasks. Alternatively, it could indicate that group level learning in this context might require other mechanisms.

And third, our study also contributes to current organizational learning studies that merge different research traditions (Argote, 2011). We provide empirical evidence of how the 'learning curve' and sociological traditions on organizational learning come together in the external context of the organization. Our research demonstrates that, at the *individual level*, learning takes place when external individuals gain experience by repeatedly performing a task assigned by the organization, *learn by doing* (i.e. learning curve). At the same time, the organization facilitates learning by providing guidelines and feedback. We consider these guidelines and feedback as instances of virtual *social interaction* between the individual and the organization. Learning occurs at the intersection between individual experience and social interaction. That is, learning is "a dynamic, two way relationship between people and the social learning systems in which they participate" (Wenger, 2000: 227). In our context, this two way relationship takes place online between people outside the organization, who individually accumulate experience; and the organization as a social system with its established rules and communication practices.

Limitations and Future research

Despite the limitation of our research being based on one case study, we believe that our insights are valid in the context of crowdsourcing as well as in other situations involving individuals connected to the organization without an employment contract. As we have explained, also in the context of services, organizations educate their consumers in performing their part of the service to achieve a good quality service experience.

Our study suggests two possibilities for future research. First, we expect that reducing knowledge variability by simplifying the task (Frei, 2006) might speed up learning, engage and retain more people. Future studies could examine whether the simplification of the task contributes to accelerated learning and retains participants, or on the contrary leads to individuals quitting due to a lack of challenge. And second, the anonymity that participants in online-based endeavours are entitled to has limited the analysis between educational background, prior experience and individual performance. New studies in a similar setting should consider including a few compulsory fields in the registration of participants, to allow such analyses without compromising privacy.

Practical implications

This study has shown that the progress made in the *Transcribe Bentham* project is mainly due to the dedication of the project staff in maintaining good relationships and clear communication with participants, and the effort of a small group of enthusiast super-transcribers that offer their time and knowledge to transcribe the complex manuscripts written by Jeremy Bentham. Though the single transcriptions submitted by hundreds of other contributors are very much valuable, our findings indicate that the project would benefit from higher quality and faster completion time if more one-time contributors were motivated to continue in the project for longer periods of time.

Project staff have focused on reducing capability variability among the participants and ensuring quality by targeting skilled people through publicity[4] in higher education and humanities specialized blogs, journals and websites, and by offering guidelines and feedback to participants. However, our study indicates that this is not sufficient to ensure higher participant retention. In line with Frei's (2006) recommendation to break the trade-off between reducing or accommodating to the capability variability of the crowd, we believe that a project could attract more people with different skills by customizing the task to fit those skills. Customization could be achieved through simplification, by dividing the manuscripts into smaller chunks of text, or by progressing through difficulty levels. We conclude that, achieving high quality contributions from external individuals and groups entails the creation of opportunities for extra-organizational learning. That is, organizations enabling experiential learning by retaining and educating them on how to perform their role or task, according to specific standards.

[4] Publicity includes the project efforts made to target participants by announcing the project in several media, and the exposure gained from press and TV reporters approaching the project. A list of the announcements and articles can be found here: http://blogs.ucl.ac.uk/transcribe-bentham/publicity/

REFERENCES

Adler, P. S., & Clark, K. B. (1991). Behind the Learning Curve: A sketch of the learning process. ***Management Science***, 37(3), 267-281.

Afuah, A., & Tucci, C. L. (2012). Crowdsourcing as a solution to distant search. ***Academy of Management Review***, 37(3), 355-375.

Argote, L. (2011). Organizational learning research: Past, present and future. ***Management Learning***, 42(4), 439-446.

Argote, L. (2013). Organizational Learning Curves: An Overview. In L. Argote (Ed.), ***Organizational Learning: Creating, Retaining and Transferring Knowledge*** (1-29). New York: Springer Science+Business Media.

Argote, L., Denomme, C., & Fuchs, E. (2011). Learning across boundaries: The effect of geographical distribution. In Easterby-Smith, M., & Lyles, M.A. (Eds.), ***Handbook of organizational learning and knowledge management*** (659 – 684). John Wiley & Sons. Retrieved from: http://kb.eblib.com/patron/FullRecord.aspx?p=693559

Argote, L., & Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. ***Organization Science***, 22(5), 1123-1137.

Berends, H., & Lammers, I. (2010). Explaining Discontinuity in Organizational Learning: A Process Analysis. ***Organization Studies***, 31(8), 1045-1068.

Brabham, D. C. (2013). Crowdsourcing: A Model for Leveraging Online Communities. In A. A. Delwiche & J. J. Henderson (Eds.), ***The participatory cultures handbook*** (120-129). New York: Routledge.

Causer, T., Tonra, J., & Wallace, V. (2012). Transcription maximized; expense minimized? Crowdsourcing and editing The Collected Works of Jeremy Bentham. ***Literary and Linguistic Computing***, 27(2), 119-137.

Causer, T., & Wallace, V. (2012). Building a volunteer community: Results and findings from Transcribe Bentham. ***Digital Humanities Quarterly***, 6(2).

Crossan, M. M., Lane, H. W., & White, R. E. (1999). An organizational learning framework: From intuition to institution. ***Academy of Management Review***, 24, 522-537.

Crossan, M. M., Maurer, C. C., & White, R. E. (2011). Reflections on the 2009 AMR Decade Award: Do We have a Theory of Organizational Learning? ***Academy of Management Review***, 36(3), 446-460.

Dyer, J. H., & Nobeoka, K. (2000). Creating and managing a high-performance knowledge-sharing network: The Toyota case. ***Strategic Management Journal***, 21(3), 345-367.

Easterby-Smith, M., Lyles, M. A., & Tsang, E. W. K. (2008). Inter-Organizational Knowledge Transfer: Current Themes and Future Prospects. ***Journal of Management Studies***, 45(4), 677-690.

Edmondson, A. C., & McManus, S. E. (2007). Methodological Fit in Management Field Research. ***Academy of Management Review***, 32(4), 1155-1179.

Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. (2011). Knowledge Collaboration in Online Communities. ***Organization Science***, 22(5), 1224-1239.

Franzoni, C., & Sauermann, H. (2014). Crowd science: The organization of scientific research in open collaborative projects. ***Research Policy***, 43(1), 1-20.

Frei, F. X. (2006). Breaking the trade-off between efficiency and service. ***Harvard Business Review***, 84(11), 92-101.

Frei, F. X. (2008). The four things a service business must get right. ***Harvard Business Review***, 86(4), 70-80.

Gibbert, M., & Ruigrok, W. (2010). The "what"and "how"of case study rigor: Three strategies based on published work. ***Organizational Research Methods***, 13(4), 710-737.

Hislop, D. (2009). ***Knowledge management in organizations.*** Oxford: Oxford University Press.

Holmqvist, M. (2003). A Dynamic Model of Intra-and Interorganizational Learning. ***Organization Studies,*** 24(1), 95-123.

Howe, J. (2006). ***Crowdsourcing: A definition.*** Retrieved from http://www.crowdsourcing.com/cs/2006/06/crowdsourcing_a.html. Accessed March 27, 2014.

Huber, G. P. (1991). Organizational learning: The contributing processes and the literatures. ***Organization Science***, 2(1), 88-115.

Ingram, P. (2005). Interorganizational Learning. In J. A. C. Baum (Ed.), ***The Blackwell Companion to Organizations.*** Blackwell Reference Online: Blackwell Publishing.

Jones, O., & Macpherson, A. (2006). Inter-Organizational Learning and Strategic Renewal in SMEs: Extending the 4I Framework. ***Long Range Planning***, 39(2), 155-175.

Knoppen, D., Sáenz, M. J., & Johnston, D. A. (2011). Innovations in a relational context: Mechanisms to connect learning processes of absorptive capacity. ***Management Learning***, 42(4), 419-438.

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. ***Organization Science***, 2(1), 71-87.

Owens, T. (2013). Digital Cultural Heritage and the Crowd. ***Curator: The Museum Journal***, 56(1), 121-130.

Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., & Kraut, R. E. (2012). Building Member Attachment in Online Communities: Applying Theories of Group Identity and Interpersonal Bonds. ***MIS Quarterly***, 36(3), 841-864.

Sauermann, H., & Franzoni, C. (2015). ***Crowd science user contribution patterns and their implications.*** Proceedings of the National Academy of Sciences.

Turner, S. F., & Rindova, V. (2012). A Balancing Act: How Organizations Pursue Consistency in Routine Functioning in the Face of Ongoing Change. ***Organization Science***, 23, 24-46.

Wenger, E. (2000). Communities of practice and social learning systems. ***Organization,*** 7(2), 225-246.

Wiggins, A., & Crowston, K. (2011). ***From Conservation to Crowdsourcing: A Typology of Citizen Science.*** Paper presented at the 44th Hawaii International Conference on System Sciences.

Zietsma, C., Winn, M., Branzei, O., & Vertinsky, I. (2002). The war of the woods: Facilitators and impediments of organizational learning processes. ***British Journal of Management,*** 13, S61-S74.