Learning with Strangers
The Value of Sets in Online Learning

Jon Dron and Terry Anderson

Technology-enhanced Knowledge Research Institute, Athabasca University, University Drive, Athabasca, Canada

Keywords: Social Media, Education, Learning Technology, Group, Network, Set, Collective, Lifelong Learning, e-Learning.

Abstract: Most research and practice relating to online and distance learning to date has focused on the social form of the intentional *group*, a named collection of people, typically hierarchically organized, with norms and/or explicit rules of conduct as well as inclusion or exclusion, membership, pacing and shared goals. The group provides a backdrop and infrastructure support for formal or informal learning activities. Since the last decade of the 20th century a different social form, the *network*, has been the subject of much research in informal and non-formal learning. Increasingly, however, we teach and we learn with and from countless anonymous others that are not formed into either identifiable networks or groups. We describe a collection of people who share little apart from interests or attributes but that none-the-less affect one another's learning as the *Set*. Under the right conditions, collective intelligence (or *collectives*) can emerge from such sets that can actively guide learning. In this paper we explore the nature of set-based learning and the role that collectives can play in helping or hindering learning.

1 INTRODUCTION

Much learning through the Internet involves following or active engagement with strangers, whether through sharing ideas and comments in blogs and websites, editing a Wikipedia page, contributing to a Q&A forum or posting to a listserv. Traditional notions of social capital, group dynamics and social contracts are significantly mutated when we are not talking with people we know or recognize, and we are in the open, away from the safety of controlled groups of people with shared purposes and norms. Beyond that, there are often emergent and/or designed effects arising from large-scale interactions that play an active role in shaping the behaviours of participants in this partly anonymous crowd. This paper is concerned with the actual and potential value of these sets of minimally connected strangers both purposefully and inadvertently helping one another to learn. As well as explaining how such sets differ from the more commonly researched social forms of groups and networks, we will be listing some of the common set tools, some of the ways they can be used for learning, some of the risks and dangers, and some potential and actual solutions to those problems.

2 GROUPS, NETS, SETS AND COLLECTIVES

2.1 The Group

The bulk of research into social learning, whether at a distance or not, has so far focused on ways that intentionally formed groups can be used to help people to learn. The *group* (or often 'team' in business circles) is a fundamental social form. It plays out in myriad ways, from the most rigid committee or court to the most informal study group or family, but it has some common features. For learning, there are familiar groups such as classes, cohorts, tutorial/seminar/working groups, teams, faculties, schools, houses and clubs. By and large they have leaders and, beyond a certain size, hierarchies of leadership. Almost all have names. All have implicit or explicit rules and rituals that govern how members should behave, how people become members and, as importantly, who to exclude.

In a learning context, most are time-limited, specify distinct goals and operate to a schedule. Groups tend to go through phases of development, such as forming, storming, norming and performing, or Salmon's five stages of e-moderation (Salmon,

2000). Groups have explicit membership: it is almost impossible to unknowingly become a member of a group and it is at least in principle possible to know the names of all the other group members. The overhead needed to organize, schedule and maintain a group is significant. Groups require commitment and do not scale well in a learning context to large numbers of people.

2.2 The Network

Over the past few decades there has been an increasing amount of research into an equally or more important social form for learning, the *network*. Every individual's networks are different from every other's, because networks are constituted of the people we know. From the weakest ties of recognition to the strongest friendships, we are normally members of many overlapping networks, often without even being particularly aware of it. Networks are mostly emergent structures based on the connections we make with others, and their edges are typically fuzzy and constantly shifting. Ideas, norms, behaviours and other forms of learning can and do spread through networks, often with amazing speed and effect. The Internet has played a major role in making networks more tangible, most notably through social networks like Facebook, LinkedIn and Google+. However, many other Internet-based systems from emails to instant messaging to blogs enable the nurturing and growth of social networks. The network is a fundamental social form for learning, described by Wenger as a community of practice (Wenger, 1998), later refined to the notion of the network of practice (Wenger et al., 2011) and providing the basis of Siemens's Connectivist model (Siemens, 2005). Indeed, networks play a crucial role in groups, connecting members within the group as well as sustaining the exchange of knowledge beyond the group.

2.3 The Set

A third important social form can also be described that extends beyond groups and networks, and that has not received anything like as much recognition in literature on learning: the *set*. **Sets are simply collections of people with shared attributes who share the same virtual or physical space.** In a learning context, the most significant shared attribute tends to be a shared interest in a topic but others may matter too, such as prior knowledge or location. In our non-virtual lives we can and do make use of sets to learn. For example, when we publish a book or a web page we normally provide categories (tags) so that people with a particular set of interests or attributes can find it. We do not know who they are but, as authors, we are communicating with and to the set of people who may find it valuable. Equally, the set can communicate with us: for instance, the fact that there is a set of people outside who are carrying umbrellas tells me that I should probably do the same when I go out. More deliberate uses of sets are common: shows of hands in a classroom, divisions of crowds by demographic, gender, or other lines are a regular feature of our lives, for better or worse.

Part of the reason for the lack of recognition of sets for learning till now is due to the fact that, in most social contexts before the advent of the mass Internet, sets performed relatively little useful work. The Internet makes it possible to interact with a vast number of people with whom we have no shared social connection at all. Much of the activity that drives Wikipedia, for instance, is from anonymous people whose only interaction is in editing one another's words. While networks and groups exist on the Wikipedia site and can play a strong role in the development of pages, there are at least as many people contributing to the site who are helping one another without ever being aware (or caring) who is helping whom. Likewise, though networks and groups exist on Q&A sites like Slashdot, Yahoo Answers or StackExchange, much of the learning that results from their use emerges from virtually anonymous interactions between people unknown to one another and not organized into groups or networks. Sets are the basis of Google Search, arguably the most significant learning technology invented in the last millennium. Sets underpin crowd-mining technologies such as Amazon's book recommendations, Netflix's movie recommendations and Pandora's music recommendations. Countless specialist sites cater for particular interests that are, by nature and our definition, set-oriented. Curation sites like Pinterest and Learnist are largely set-oriented, focusing on topics rather than communities. Twitter hashtags are primarily concerned with sets, not networks or groups. Usenet newsgroups and email listservs have long been an important source of knowledge and dialogue, often among strangers sharing nothing but an interest in a topic or need for topic-specific information. Despite the popularity of group-supporting tools like learning management systems and network-nurturing tools like Facebook, Academia.edu and LinkedIn, set-based interactions are the dominant social form in Internet-based

independent learning and may soon be in formal education as well.

2.4 The Collective

Set-oriented systems can be wild places, full of half-truths and falsehoods as much as rich and meaningful information, not to mention abusive, malevolent and mischievous contributions. This is overcome in part through reification of the conversation, so that individuals can choose the most compelling solutions and arguments. More significantly, almost all successful systems of this nature incorporate crowd-sourced algorithmically collated metadata like ratings, likes, reputation measurement and filtering tools so that the crowd can collectively guide its own members.

We refer to the outcome of this algorithmic combination as a *collective,* using the term much like the creators of Star Trek's Borg to signify a single entity made up of many independent entities acting as one. Collectives combine the behaviours of many people through one or more algorithms in order to provide help, guidance or structure to otherwise overwhelming or ambiguous content generated by the crowd. The algorithms may be provided by machines, such as in rating systems, or collaborative filters, or by people, such as when people are collectively drawn to active sites or repelled from those that are too active, or both, as we see in people's reactions to the search order of a Google search or the weightings of tags in a tag cloud. In many cases, processing is split between a machine and the heads or hearts of human beings, the machine offering alternatives according to one set of algorithms and people making choices using others. A collective is not a social form as such, but the emergent result of people interacting, directly or indirectly, with one another.

2.5 Set Combinations

Social forms seldom exist in isolation. Sets may be a supplement or a pre- or post-emergent form of traditional group-based learning, existing networks or conventional individual study. Equally, the social forms we describe are not binary categories but are more like primary colours that often occur in blends. For example, at our own Athabasca University, our individualized study model means that students are self-paced, choosing when and how they work over a six-month period. It is thus rare for two students to be working on the same things at the same time. Despite this, forums other social tools are normally

provided for each course. Although courses share some group-like features including rules, shared goals and hierarchies, students do not form teams, seldom know others, do not collaborate and are not expected to work together. Their interactions are thus notably set-like. What they share helps them to solve problems, alleviate a sense of isolation, and discover different ways of seeing a subject. Many large MOOCs, though they may have designs that resemble those of conventional group-based university courses, are more set-like in social form, for similar reasons.

3 WHY DO PEOPLE CONTRIBUTE TO SETS?

For many contributors to the public good, social capital plays an important role: by providing help to others, one is increasing one's own social capital, with consequent gains for all concerned (Nemoto et al., 2011). This is equally true in a learning context (Daniel et al., 2003). However, this is not the whole story, even in tight-knit social networks, where expectations of reciprocity may not play a dominant role (Wasko and Faraj, 2005). A survey of frequent contributors to Wikipedia found that five of the top 67 editors (those who have made at least 500 edits) were known only by their IP addresses (Various, 2005). Amongst these anonymous contributors there can be no expectation of reciprocal social benefits. As a species, we have an evolved tendency towards altruism that cannot be simply explained away by assumptions that people rationally weigh costs and benefits. We are genetically inclined to help one another (Wilson, 2012). Beyond anonymous contributions, many sets emerge as a side-effect of other interactions. For example, academics may publish blog posts primarily for the benefit of a small subset of people in their own networks or groups, while knowing that there is an added benefit that their writings might be read by the set of others with a similar interest.

4 SET LEARNING

Set-based learning tends to be appropriate when the objectives of learning are already known. It is well suited to information seeking, inquiry-based and problem-based pedagogies, where goals are known and the learner already has some subject knowledge.

While people often help one another in sets,

there tends to be little or no deliberate collaboration because there are few opportunities for sustained interaction, no shared projects, and limited scheduling of activities. Conversely, there are also minimal temporal or spatial constrictions on independent learning. Cooperation (not collaboration) is the dominant form of working together, in which learners working individually contribute to the learning of others. There is more sharing with others rather than direct dialogue and, when dialogue occurs, it tends to be fleeting and limited in scope. Where coordination does occur, it is either through centralized methods like FAQs compiled by individuals, or more sophisticated structural processes such as the forking process used in Github, that enables people working independently to contribute to one another's work.

Apart from sets that form around temporal events, most sets tend to eschew schedules and pacing. People tend to contribute as and when they want or need to do so. For those seeking answers to problems or discussions about issues, this can be frustrating, unless the set is sufficiently large to ensure a constant succession of contributors. However, the almost ubiquitous reification of previous interactions (including recommendations) means that answers previously given at one moment can continue to provide value to later-arriving members of the set.

Sets have great value in forming and building learning networks and even groups. For instance, on sites that form around (say) support for a specific piece of software, there is typically a caucus of enthusiastic contributors who come to know and respect or at least recognize the strengths and limitations of one another, leading to what may often be rightly described as a community. While there may be hundreds or thousands of occasional contributors in such spaces, and countless people who do not contribute, but do read, such spaces often contain rich social networks as well as sets. The non-contributors in such set-oriented spaces are often misleadingly referred to as 'lurkers'. This is a consequence of failing to recognize that sets are not communities as such, mostly lacking the norms and network bonds that hold communities together. It is as meaningless to describe readers of books as lurkers as to describe members of sets that way.

Sets are typically great for finding diverse views and perspectives, inasmuch as the shared attributes that bind the set together may have little to do with any other shared values. There are typically few dangers of group-think, nor of only connecting in yet another network with like-minded people. Despite

the potential for this diversity they may also reveal underlying homogeneity that can be used as a basis of more intensive interaction. It should also be noted that some shared attributes such as religious belief, occupation or cultural origins, may be a shorthand for a cluster of shared attributes or set memberships. There is a world of difference between the set of religious fundamentalists and the set of people interested in learning to sail.

5 SET DISADVANTAGES

5.1 Focus

In order to learn in a set it is normally necessary to know what one wishes to learn. Unfortunately, knowing *that* is one of the most common challenges faced by a learner. Until one has been immersed in a subject, it is hard to know what questions to ask, - what sets to align with. There are some solutions. Many Q&A forums, for instance, are divided into categories such as 'help for beginners' and 'advanced topics', creating subsets with a learning focus. Similarly, every Wikipedia page supports and is the focus of a different set. Wikipedia provides plentiful links within each page to other pages, that a learner can follow in order to gain a grounding in a topic as well as to get foundational knowledge in many areas – exhibiting the learning potential of the set. However, set-based learning can be overwhelming unless the learner already knows the information he or she needs to seek. The paths through potential answers are multitudinous, so set-based learning can be circuitous and inefficient. Moreover, the information that is available may often be contradictory, and it can be hard for a beginner to distinguish the good from the bad.

5.2 Depth

Related to problems of focus, set-based learning typically tends to involve brief exchanges rather than sustained dialogue. This is fine if one needs an answer to a programming problem, but not great if one is seeking to become a medical doctor, where lengthy study crossing many disciplinary boundaries may be needed, and where a sustained path may need to be planned, with dependencies and prerequisites at every stage. The set may be able to provide help with constructing or advising on such a path, but it requires a fair degree of self-discipline, independence and self-determination to succeed. Typically, sets may provide help and support but, for

longer learning journeys, are often best supplemented by networks and/or groups. Sets can provide the seed for these to emerge, with phases of peripheral participation leading to stronger involvement with networks of learning partners as time progresses. However, for set-oriented approaches like xMOOCs (sets with an interest in subject X) that often use group-oriented methods like tight schedules, there may be challenges of insufficient time for networks to form and learners needing group support may be set adrift.

5.3 Trust

One of the biggest problems faced by set-based learners is that anonymity makes it more likely that there will be trolls, spammers, scammers and other undesirables. Even when intentions are good, sets often contain members with limited knowledge as well as those with too much knowledge, whose attempts to help may be positively harmful. Inaccurate or scanty knowledge may result in poor foundations or wasted work, while excessive complexity or jargon can be demotivating to someone trying to make sense of the basics. Division of set into subsets with greater focus can help here, as can enthusiastic moderators, but more complex collective tools are often needed to address this problem.

5.4 Diversity

Part of the value of sets lies in the diversity of opinions, skills and interests of set members. However, this can come at a high price because people with different cultures, different vocabularies and different understandings may cause confusion, upset one another, or fail to communicate effectively. Sets are fertile ground for flame wars, angry debate and what some set members will see as irrelevant or unimportant. At best this can be inefficient, at worst it will drive people away from the set. Thus, this strength of diversity is also a potential weakness of disharmony.

6 THE ROLE OF COLLECTIVES IN SETS

Given the aforementioned difficulties, learning within sets can be frustrating, misleading, circuitous and poor for motivation. Collectives can provide the missing pieces to replace some of the guidance roles of the teacher and can make up for the lack of personal connection and relatedness that occurs in networks. The general principle behind any collective is that the actions of many people are combined, processed and represented, typically as recommendations, or for filtering, or to structure information or to suggest a path through it. Collectives can filter and help make sense of the information generated by the set (and, to a lesser extent, the net and the group). For example:

- An automated collaborative filter can find others with similar patterns of interest or behaviour, and recommend content that may be of value.
- A tag cloud can show topics of interest to a set, helping to get a better sense of the overall shape of a subject area and to make it easier to know what to look for, suggesting other things that may be of interest.
- A reputation system can identify individuals who have been found to be trustworthy or knowledgeable within a subject area.
- A rating system can help promote good answers/solutions/recommendations and demote bad ones.
- A data visualization tool can graphically display activities, actions or ideas of a set of learners.
- A crowd-sourced spam filter can help to remove content that is injurious or irrelevant

Collectives based on sets *can* be an embodiment of the wisdom of the crowd, with relatively few of the problems that can arise when individuals are connected or know what decisions others are making (Surowiecki, 2004). Sometimes, sets of moderately informed people can outclass experts when dealing with a range of tasks (Page, 2008). Collectives are, however, only as smart as the algorithms that underlie them and the combined wisdom of the crowds that feed them. This means that they tend to be susceptible to some common flaws, including:

- The Matthew Effect (Merton, 1968), in which the rich get richer and the poor get poorer, an out-of-control path dependency that makes it hard for better novel solutions to gain a foothold and the rewards priority and familiarity more than quality.
- Filter bubbles (Pariser, 2011), in which we tend to see things that resemble what we have already seen, limiting opportunities for serendipity and discovery of novelty. This is especially risky for learners who, by definition, need to enter novel territory.
- Lack of pedagogical model, so that it is not

always value to learners or even learning that is valorized in the results. Relatively few collectives explicitly support learning and most rely on some variation of popularity or commonality measures albeit, in the case of more sophisticated tools like collaborative filters, with significant personalization.

- Intentional abuse, in which mischievous or malevolent people, especially when working as in consort, can subvert or overly influence a system. 'Google bombing' and search-engine optimization strategies are good examples of this.

- Selection bias, in which a distinctive subset of individuals provides a biased collection of raw data on which to operate. For example, a student or a set of experts may fail to consider solutions to problems that are unconventional, and so miss some important opportunities.

While collectives have been used to good effect in an educational setting as well as offering a lot of value to informal and non-formal learners, it remains an important research area to find ways of adapting them effectively to the distinctive needs of learners.

7 CONCLUSIONS

The set is an under-researched social grouping that we have only recently begun to explore. The set has increasing importance as we move away from the familiar formal learning approaches of institutions that worked well in an industrial face-to-face context but that do not operate so well at Internet scales, and that do not cater well for informal or just-in-time learning. As well as being crucial in supporting day-to-day lifelong learning, the social form of the set dominates in large-scale MOOCs. However, many MOOCs are designed as though they were groups of a conventional academic variety, with schedules that assume group-like engagement and commitment, discussion forums that are often over-populated, fuzzy in purpose or that assume collaborative rather than cooperative pedagogies. As a result, they often carry unrealistic expectations of trust and shared intent that, in a large and diverse population, are unlikely to be achieved. This paper has begun to scratch the surface of how and why we might use sets for learning, as well as some of the pitfalls that await the unwary. We continue to research the differences and to build tools to support sets for learning. In our forthcoming book, Teaching Crowds (Dron and Anderson, in press), we explore these issues in greater detail.

REFERENCES

Daniel, B., Schwier, R. & McCalla, G. (2003) Social capital in virtual learning communities and distributed communities of practice. *Canadian Journal of Learning and Technology.*, 29,

Dron, J. & Anderson, T. (in press) *Teaching crowds: social media and distance learning*. AU Press, Athabasca.

Merton, R. K. (1968) The Matthew effect in science. *Science*, 159, 56-63.

Nemoto, K., Gloor, P. & Laubacher, R. (2011) Social capital increases efficiency of collaboration among Wikipedia editors. *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*, 231-240.

Page, S. E. (2008) *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies (New Edition)*. Princeton University Press,

Pariser, E. (2011) *The filter bubble : what the Internet is hiding from you*. Penguin, New York.

Salmon, G. (2000) *E-moderating: The Key to Teaching and Learning Online*. Kogan Page, London.

Siemens, G. (2005) Connectivism: a learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2,

Surowiecki, J. (2004) *The Wisdom of Crowds*. Little, Brown, London.

Various (2005) What do Wikipedians do all day? 2013, retrieved from http://en.wikipedia.org/wiki/User: Statistics#Case_1:_Anon_Surprise.21 12th Februar 2014.

Wasko, M. M. & Faraj, S. (2005) Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice. *MIS Quarterly*, 29, 35-57.

Wenger, E. (1998) *Communities of Practice: Learning, Meaning and Identity*. Cambridge University Press, New York.

Wenger, E., Trayner, B. & de Laat, M. (2011) Promoting and assessing value creation in communities and networks: A conceptual framework. *The Netherlands: Ruud de Moor Centrum*,

Wilson, E. O. (2012) *The social conquest of earth*. Liveright Pub. Corporation, New York.