New Perspectives on Great Questions: What can MOOCs Tell Us?

David Robertshaw [D.Robertshaw@derby.ac.uk], School of Nursing and Professional Practice, Ivana Babicova, College of Health and Social Care, University of Derby, Kedleston Road, Derby, DE22 1GB, United Kingdom

Abstract

The aim of this paper is to consider the role of massive open online courses (MOOCs) as a source of rich natural language data to be mined and analysed, thereby providing new perspectives and insights to questions, theories or experiences in relation to health and healthcare. It will consider the concept of "great questions" and why these are challenging to understand but suggests that MOOCs may enable us to discover some answers. It considers existing practice, emerging concepts and recommends future directions in this field.

Keywords: MOOCs, research, data mining, experiences, health

New perspectives on great questions: what can MOOCs tell us?

This paper considers the role of massive open online courses (MOOCs) as a source of rich natural language data to be mined and analysed, providing new perspectives and insights to questions, theories or experiences in health and healthcare. It will consider the concept of "great questions" and why these are challenging to understand but suggests that MOOCs may enable us to discover some answers by collecting and analysing natural language data found in MOOCs. It considers existing practice, emerging concepts and recommends future directions in this field.

Massive open online courses (MOOCs) are an educational phenomenon, with the first course beginning in 2008: "Constructivism and Connective Knowledge". Shortly thereafter, a free online course entitled "Introduction to Artificial Intelligence" attracted over 160,000 learners (Rosewell & Jansen, 2014). The underlying concepts of MOOCs and how they operate are not new. This type of learning has roots in distance and distributed learning, but MOOCs have a basis in Connectivism, an educational theory which integrates principles of chaos, network, complexity and self-organisation theories (Siemens, 2005). MOOCs are an application of network learningSince 2008, over 900 Universities have delivered more than 11,000 MOOCs to over 100 million students, although student retention and completion rates are low and there are signs that course registrations are decreasing rather than increasing with 20 million students taking MOOCs during 2018 (Shah, 2018). Completion rates may be low despite increasing participation because students often use MOOCs for casual learning, "dipping in" to topics which interest users or that are important to them at that time. MOOCs are large, open to all, online and accessible through the internet and may exist as a programme or course. Beginning as part of the Open Educational Resource movement, MOOCs attract many thousands of students. Openness relates to the use of open-source software, open curricula, open sources of information and open assessment processes (Rodriguez, 2012). MOOCs are massive and some of the largest courses have had up to 300,000 participants. They have become a popular way of learning, making University-level education accessible to all. Learners can study MOOCs on a wide range of topics and each has a different style or approach which may be dependent upon the author, the institution, the platform or the topic (Reich & Ruiperez-Valiente, 2019). MOOCs are a mode of online learning. Online learning, also known as computer-based instruction, involves learning undertaken using a virtual learning environment (Knight & Price, 2016). MOOCs may adopt a variety of different models and designs, and therefore not all features of MOOCs apply to all MOOCs. For example, the openness of MOOCs has in many ways decreased in recent years with Universities, companies, and other education institutions delivering MOOCs which are less open with more restrictions. Some MOOCs do not run in closed environments, and may not specifically run on a learning platform.

Although MOOCs allow learners to expand their knowledge and understanding, they also provide unique opportunities for pedagogic and educational researchers to innovate in teaching and learning methods. New pedagogic ideas and/or methods can be tested at scale with many users, with learners taking alternative routes to allow comparison of learning activities. This testing of new ideas feels safer because it is happening outside of the traditionally formal and regulated education environment, although some MOOCs are regulated. Results from these insights could then be applied in the classroom or online learning environment. MOOCs also provide opportunities for researchers to ask questions about social issues or concepts in order to gain new insights into the nature and content of responses. This is a recent area of interest for researchers and uses approaches common with Big Data analysis. MOOC datasets are

large and rich, allowing analyses not previously possible. Recent examples of these analyses have been on important social issues like dementia (Robertshaw & Cross, 2019a), integrated care systems (Robertshaw & Cross, 2019b), palliative care (McInerny et al., 2018), and end of life care (Tribett et al., 2016; Rawlings et al., 2017).

Historic approaches to great questions

The data included in MOOCs gives us opportunities to answer great questions. A *great question* in terms of this discussion is one that has multi-dimensional meaning with a variety of concepts, understanding and perspectives. Some examples of great questions might be "what is the meaning of life?", "why was I born?", "who am I?", "why is there suffering?", or "why am I poorly?" It is not clear who decides what a great question is, rather researchers may identify questions related to their courses or areas of interest for researchers and the field. Questions could vary from those existential ones above to more simple ones like "do MOOCs enable learning?", and "are participants participating?" These questions often do not have a single answer and stimulate debate. They are also a core part of the fabric of our being. They are part of the human condition and many philosophers and researchers have attempted to answer them: this is the purpose of research itself. Humans have written and thought extensively about these types of great questions throughout the ages. Initially, this began as oral histories and storytelling but progressed quickly to writing and typing. Researchers may take a social science approach to understand these issues and one such approach is examining the written word to bring new meaning. As long as there has been writing and text, people have attempted to analyse it to search for new, alternate or hidden meaning. Text has been analysed since the 1200s when friars began to produce indices of the Bible, which people could use to compare and contrast verses. These ideas and methods continued exponentially and text is analysed every day by governments, corporations and researchers to bring meaningful insights to a variety of contexts. The applications are vast: from intelligence gathering to awareness and prediction of market trends to tracking people, illness and disease. Natural language processing is the application of technical and computational analysis and synthesis of language, bringing together the fields of linguistics and computer science to process language data. Natural language processing has brought new insights into large volumes of text, and this is a fairly recent area of research in health and healthcare. When combined with data mining (where data is scraped, manipulated and analysed), natural language processing has provided a powerful interdisciplinary platform through which we can collect (mine) data, and analyse it (process) with pre-defined tools and techniques to bring new insights to difficult questions (Ignatow & Mihalcea, 2017). Although this field is established in the realms of data science, its use is yet to be fully explored in the health and social care context.

Historic approaches to answering questions or bringing new meaning/understanding have taken a range of approaches. Often the study of experience, attitude or perception is through small-scale studies. Whilst these results may be reliable and valid in that context, they may not be generalizable to the population at large. However, it is not always practical to conduct large-scale studies. Although there are many examples of these (for example, the Northern Irish Life and Times survey, Censuses), they are usually expensive and can often only be conducted by publicly funded bodies. This can prevent researchers from conducting large-scale studies. Due to feasibility and financial reasons, they may instead wish to conduct small-scale studies or phenomenological studies including a very small number of participants. These phenomenological studies do not lack rigour, but they may lack generalisability, validity and reliability. This is acceptable for phenomenological studies; however, and there are advantages to finding out the attitudes, perceptions and experiences of larger populations. There is often value in understanding how individuals feel and perceive certain aspects of their lives. These individual differences are a crucial part of how our societies and communities are formed and maintained. One of the problems with this type of study is that "experiences", which are embodied and lived through sensory and affective modes (Pink et al, 2016), vary between individuals and they are highly subjective. Experiences are often difficult to articulate or accurately characterise. Therefore, it has historically been very difficult to seek answers to big questions, especially those which relate to our health, and experiences of healthcare problems.

What questions could be answered?

Health is generally regarded as one of the central priorities of our lives. Phenomenological studies have characterised experiences of health conditions for example breast cancer (Archer, Holland, & Montague, 2016; Holland, 2016; Holland, Archer, & Montague, 2014) and healthy eating (Holland, Peterson, & Archer 2018), however, there is little data examining how the public at large perceives health conditions. Small scale studies have been conducted manually to analyse small volumes of data, but analysis of MOOC data presents opportunities for characterisation and generalisation. The applications of this understanding extend to government policy, medical research and clinical evidence-based practice.

For health researchers, these questions are important. Experiences can be categorised and characterised, with needs, requirements and interventions identified. For example, a person living with hip osteoarthritis may experience pain and discomfort. It can then be identified they need pain relief, a hip replacement and physiotherapy. Although much of our health research focuses on treating disease and illnesses, it is equally as important to understand the experience of it. For existing patients, we can use their electronic patient record to analyse notes and observations about their care. However, the electronic patient record is written by a bystander rather than the patient themselves: any observation is, therefore, a proxy or subjective

perception of their real experience. For example de Ruddere et al's (2011) study found that likeability affects the likelihood of receiving analgesic administration: if the observer likes the person they are observing, they are more likely to prescribe painkillers. Their observation was altered by their subjectivity. Many authors believe that all research (regardless of approach) includes an element of subjective perception, and the approach of researchers should be to identify "subtle realism": seeing research as an attempt to represent reality through a particular viewpoint rather than to attain the "truth" (Hammersley, 2013; Kirk & Miller, 1986; Mays & Pope, 2000). Researchers can use bracketing, based on the philosophy of Kant, to suspend judgement about the "real" world focusing on experience seeing this as a phenomenon rather than reality (Gearing, 2004; Gelling, 2010; Koch & Harrington, 1998).

Researchers are now using large volumes of data from social media networks and online social environments to answer great questions. Online participant metadata from MOOCs, social media and other online fora present some of the largest and richest datasets currently available. A big data source, participant metadata and contributions in the form of natural language data are vast with great variety. This information may allow discovery of new insights into problems that present challenges currently unsolvable with manual analysis, or to discover questions not yet asked. Twitter, for example (an environment similar to MOOCs), has over 326 million users per month who tweet 500 million times per day (Cooper, 2019). Each tweet consists of up to 280 characters. The amount of text produced each day is therefore up to 140,000,000,000 characters. Twitter can be used in a similar way to MOOCs and twitter has been used to explore a variety of health conditions and public health trends including influenza (Chew & Eysenbach, 2010), e-cigarettes (Huang et al, 2014), cardiovascular disease (*Sinnenberg et al., 2016>*) *and mental health (Coppersmith, Dredze, & Harman, 2014).*

Towards a theoretical approach

MOOCs can be characterised as social worlds. Social worlds, an abstract concept, are bounded communities allowing exploration of open-ended questions (Pink et al, 2017). These social worlds can be explored with digital ethnography approaches. Digital ethnography, sometimes known as netnography can be described as a methodological approach examining people's interactions and contributions in the online environment (Kozinets, 2009). Netnography as a methodological approach may provide limited access to true participant identity and demographic information, but Kozinets (2002) sees this as a necessary and acceptable shift from traditional ethnographic approaches. Studies in this paradigm often use an inductive logic approach, where researchers begin by analysing the data then allowing conclusions to emerge organically from the analysis (Ignatow & Mihalcea, 2017).

Brooker et al. (2016), when referring to the analysis of social media content, see the understanding of research phenomena as co-constructed results of interactions. They characterise the results of studies in this field as being bound to a particular "assemblage" of components, which make the outcomes or conclusions specific to that context and that context alone. The conclusions of a study in this paradigm are a result of the varying components coming together: the user, the researcher, the medium, and the technology. A suitable comparison for this might be an oil painting: a variety of components must come together to make the "experience" including the canvas, the frame, the paint, the painter and the viewer. The experience is only a result of those components coming together and means that any conclusions or outcomes drawn from a particular context may only be applied to that context. We may make generalisations or comparisons, however, due to the nature of the phenomena, these may, ultimately, not be relevant or useful. The reality seen within the assemblage is a representation of the real reality, which may no longer exist or even have existed at all.

"Is it ethical?"

For research activities, participant consent is usually required to be able to use an individual's data. Generally, "conventional" research studies can gain this easily by asking participants to review an information sheet and then confirming their informed consent. Online social worlds such as MOOCs present some additional ethical dilemmas. For example, participants may not disclose their true identity and may have alternative personas which they portray only in the online community. Conway (2014) suggested some of the ethical issues may include privacy and gaining informed consent. Ethical principles and research governance committees may not yet have caught up with current practices in the digital world (Tattersall & Grant, 2016), and therefore may require some adaptation as social attitudes change. Companies are regularly using online data and participant metadata for commercial applications, and governments are using the same data for surveillance. In some applications, the use of this data may be ethical even without consent for specific purposes (for example, Coppersmith, Dredze, & Harman, 2014) however there are opportunities for nefarious usage of this data as has been seen, for example, in the case of Cambridge Analytica (Ingram, 2018).

There are greater ethical concerns with using participant metadata in textual analysis, whether this originates from MOOCs or other sources. MOOC participants are usually accessible and consent can be obtained with the standard methods, however, in some other situations (for example, Twitter), it may not be possible to obtain informed consent due to lack of access to users or impracticality. For example, it is difficult to gain consent from millions of users. There is an interesting interplay in the Twitter context between regulations about privacy versus the expectation of users for their information and content to

have public exposure (Sveningsson, 2003). Sudweeks and Rafaeli (1996) believed that social media data is in the public domain and is therefore accessible for research purposes. This is influenced by whether the domain is open or closed, and there is a greater expectation of privacy in a closed domain. Interestingly, Twitter's privacy policy (Twitter, 2018a), developer agreement and developer policy (Twitter, 2018b) adopt an open approach, expressly permitting the use of its data for analysis through its application programming interface (API). MOOC privacy policies vary but generally, they are specific about the provider's use of their data for analytical purposes.

"Should I be doing it?"

Whilst it is possible to access data from a wide range of sources, using that data is in an ethical minefield, there is still the question of whether a researcher is allowed to use the data for research purposes. Fundamentally, the General Data Protection Regulations (GDPR) govern how data may be used within the European Union. This regulation requires consent by contributors for researchers to use their data. In some circumstances (for example Twitter data) there is a contractual permit to use the data for specific purposes. Social media data is generally regarded as being in the public domain and is therefore available for research. This approach has been advocated recently in studies conducted by Scanfeld, Scanfeld, and Larson (2010) and Reavley and Pilkington (2014), where Twitter is the source data. Twitter's privacy policy (Twitter, 2018a), developer agreement and developer policy (Twitter, 2018b) expressly permit the use of its data for analysis through its application programming interface (API). Twitter's privacy policy states: "Twitter is public and Tweets are immediately viewable and searchable by anyone around the world." (Twitter, 2018a). This policy is contradictory to the British Psychological Society's Code for Internet-Mediated Research (British Psychological Society, 2017) which requires consent to be provided by every participant. MOOCs operate on a much smaller scale, and in closed environments, where consent is much more easily obtainable. Consent in MOOC circumstances, if given, should be obtained through a normal agreement process unless expressly included in terms of service, terms and conditions or another similar document which users must agree to before using the service. It must be clear what the purpose of the data processing is, and the problem is this is not always clear even to researchers at the beginning of their research. This also translates to debrief, as it is not possible to debrief participants if their identity is not known. This is contradictory to human research conventions.

"Is it true?"

One of the main issues with online social worlds is that the concept of identity (including names, gender, age and other demographics) is a fluid concept. For example, identities may either be true (such as is required by some social media platforms), mostly true (where people alter details such as their age but mostly remain true to their real identity), or false (where the identity is completely fabricated). Participants may choose to be real or false out of personal preference, or because of nefarious, malicious or disruptive reasons. It is important to consider whether these are issues or not: fabricated online identities could alternatively be seen by participants as being their true identities; they may believe their real-world identities are actually the fabricated ones. Therefore, it is difficult to determine the nature of whether anyone really has a true identity or not. Within this context, it is acceptable to consider that participant metadata relates to some sort of person, whether real or not real, who might actually be the same person. This same person could, however, manage multiple accounts on the social media platform to influence or dominate the argument. Some of these issues have less relevance and impact on the data or meaning drawn from it; for example, gender as a concept is much more fluent and is not questioned or considered to have as much relevance as it once did.

There is also the issue of who is the recipient and analyser of the data. Analysis of text has historically been performed by people, however increasingly artificial intelligence and machine learning are performing analysis independent of people. Eventually it will not be clear why analysis has been performed in certain ways or how the outcomes of analysis have been arrived at: arguably this happens already with people, but our increasing reliance upon computers could mean that data is not real or true, does not relate to real or true people and is not being analysed in any true, real or meaningful way.

These issues are resolvable within the context of the assemblage: the results are valid to that specific context only. This is difficult, however, as researchers wish to make generalisations and conclusions about their work which are applicable to practice. This problem, integrating theory with practice, is not a new issue. Practitioners should be cognizant of research outcomes and apply them to their own context, within the confines of assessed risk as with any other research. Demographic information, as is the case for any other type of research or from any other source, may be unreliable: if participants wish to give false information regarding their identity they can. Arguably unless the research needs to examine specific demographic information, more often than not this should not hinder the research results because the views, opinions and experiences are more valuable than demographic information.

What have MOOCs told us?

Using MOOCs to bring new insights is a reasonably new methodological approach which has promise. For example, some research has been conducted with MOOCs in the health arena. Robertshaw and Cross

(2019a) conducted research into the experiences of integrated care for dementia. Their study found participants believed integrated care should be person-centred and holistic, involve a multidisciplinary team of health and social care practitioners, and involve the person, their family and the wider community. Their study also found that integrated care was viewed positively. Robertshaw and Cross (2019b) also did an analysis of the roles and responsibilities of people working in dementia care and found there were boundaries and issues, identifying some enablers to promote integrated care: communication technologies, shared care records, care planning and education. They also identified barriers including funding, role conflicts, time constraints and time-consuming paperwork. McInerney et al. (2018), in their study using a MOOC on palliative care, found that there was a lack of palliative care literacy and that people were seeking palliative care and support but there was a lack of information available that was specific to their context. Rawlings et al. (2017) also used a MOOC on palliative care to discover the nomenclature used to describe death and dying: they found a diversity of euphemisms were used, and participants offered commentary on their purpose and use. Their findings showed that people have become creative in their use of euphemisms. Tribbet et al. (2016) in another palliative care MOOC, also demonstrated increases in awareness, knowledge and understanding of palliative care. Annear et al. (2016) used a MOOC to validate a measurement scale to assess dementia knowledge: their study compared the new scale with an existing one, and tested it with many thousands of participants.

Each of these studies was large scale, used MOOCs as the platform, and participants were from around the world. The course designers used MOOCs to ask questions, seek answers, and to examine the way people talk about and discuss topics of importance to them: care, dementia, death, dying, and knowledge. These studies have brought new insight and awareness into the topics they have examined. Often, some of the findings were already known but large-scale research with MOOCs has validated previous findings from small scale studies.

Future directions

Although MOOCs are already being used for this purpose, research within MOOCs is often accidental or a by-product of the volume of data collected. It would be helpful to, at the design stage, include research questions or experiments which form part of the overall MOOC construct. Intentional research design would allow researchers to ask specific questions aimed at answering important questions, which otherwise may not be possible to ask in other ways. For example, Robertshaw and Kotera (2019) applied the Northern Ireland Life and Times questionnaire on dementia to reach a wider audience, using the tool to test the impact of their MOOC on dementia.

A major issue of research in this context is ethical approval. Online research is still an emerging field and ethics committees may not be well-versed in their methodologies and approaches. There are helpful textbooks, articles and guidelines to assist decision making but some of the issues around consent, in particular, are difficult to navigate. Software platforms have specific terms and conditions which may be useful to understand the legal context of MOOC-related research, however this area is under-developed. In particular, one area of challenge is reporting of participant demographics, which may not be permitted depending on the software or tool. If users do not give permission for this to be shared when signing up, this is not a useable or reportable data source.

This research with MOOCs highlights an approach which could be expanded further into other areas: there are vast online social worlds with rich natural language content to be explored. One such area is Twitter, for example, where there are enormous volumes of data. This data is already being used for commercial purposes but is yet to be fully explored for health research. There are some examples of Twitter-based research in the health field (Robillard et al., 2013; Cheng, Liu, & Woo, 2018; Oscar et al., 2017), but this are requires further development. The next step for this general methodology of using MOOCs for insights is to push beyond what we already know, to find new knowledge, by finding those online places where participants are discussing the topic of emerging interest as it happens. This could provide researchers with real-time information about the perceptions, opinions and feelings of participants towards a particular issue. For example, Twitter has been used to track deaths from a disease, but it could also be used to characterise their experience of dying by analysing tweets related to this topic. This analysis may bring new insights in to the experience of dying, which might not otherwise be available.

MOOC participant metadata is currently being used on a surface level to track participation, test ideas or be analysed for new insights. But this data has greater potential, yet to be unlocked, to bring deeper levels of insights about participant behaviours, thoughts and feelings although the reliability of this information may begin to decrease as researchers draw conclusions about it.

Conclusion

The use of participant metadata from online social worlds is a controversial methodology and there are many complex issues regarding its use in health research. MOOCs have the potential, as a regulated and reasonably reliable online world, to be used as platforms for research. Small groups of researchers are already conducting this research for a limited range of topics in healthcare and other areas, but there is potential to expand their use to other areas of practice and research. MOOCs have already brought new

insights to a range of health topics, but their potential is still to be fully explored in relation to a broader range of insights.

MOOCs are continuing to be popular among users, and their use is likely to increase. They provide a rich and diverse source of research participants who are able to participate in research studies whilst also learning and developing new skills and knowledge. This area of research is promising in healthcare, where MOOCs can provide new insights into healthcare conditions. There remains, however, many challenges relating to the use of MOOCs and online communities like them as a source of natural language data. Specifically, ethical issues, identity issues and issues of validity and reliability remain. It is often not possible to generalise conclusions or findings and relate these to other areas of practice however MOOCs could be used to triangulate findings from other studies, or perform an initial feasibility study function.

The aim of this paper was to consider whether MOOCs have a role as a source of natural language data. MOOCs are a good source of natural language data to be mined and analysed, however this should be considered within the context of whether the data is real, whether the participants are real, if the analyser is real and ultimately whether there is such a thing as reality.

References

1. Annear, M. J., Eccleston, C. E., McInerney, F. J., Elliott, K. E. J., Toye, C. M., Tranter, B. K., & Robinson, A. L. (2016). A new standard in dementia knowledge measurement: comparative validation of the dementia knowledge assessment scale and the Alzheimer's disease knowledge scale. Journal of the American Geriatrics Society, 64(6), 1329-1334.
2. Archer, S., Holland, F. G., & Montague, J. (2018). 'Do you mean I'm not whole?': Exploring the role of support in women's experiences of mastectomy without reconstruction. Journal of Health Psychology, 23(12), 1598-1609.
3. British Psychological Society (2017). Ethics Guidelines for Internet-mediated Research. INF206/04.2017. Leicester. Retrieved from https://www.bps.org.uk/sites/www.bps.org.uk/files/Policy/Policy%20-%20Files/Ethics%20Guidelines%20for%20Internet-mediated%20Research%20%282017%29.pdf
4. Brooker, P., Barnett, J., & Cribbin, T. (2016). Doing social media analytics. Big Data & Society, 3(2), 1-12.
5. Cheng, T. Y. M., Liu, L., & Woo, B. K. (2018). Analyzing Twitter as a Platform for Alzheimer-Related Dementia Awareness: Thematic Analyses of Tweets. JMIR Aging, 1(2), e11542.
6. Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PloS one, 5(11), e14118.
7. Christensen, C. M. (2013). The innovator's dilemma: when new technologies cause great firms to fail. Harvard Business Review Press.
8. Conole, G. (2016). MOOCs as disruptive technologies: strategies for enhancing the learner experience and quality of MOOCs. Revista de Educación a Distancia, 50, 1-18.
9. Conway, M. (2014). Ethical issues in using Twitter for public health surveillance and research: developing a taxonomy of ethical concepts from the research literature. Journal of Medical Internet Research, 16(12), e290.
10. Cooper, P. (2019, October 30). 28 Twitter Statistics All Marketers Need to Know in 2019. Hootsuite [Blog post]. Retrieved from https://blog.hootsuite.com/twitter-statistics/
11. Coppersmith, G., Dredze, M., & Harman, C. (2014). Quantifying mental health signals in Twitter. In Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality. Workshop. Baltimore, USA.
12. Gearing, R. E. (2004). Bracketing in research: A typology. Qualitative health research, 14(10), 1429-1452.
13. Gelling, L. (2010). Phenomenology: the methodological minefield. Nurse researcher, 17(2), 4-7.
14. Goldberg, L. R., Bell, E., King, C., O'Mara, C., McInerney, F., Robinson, A., & Vickers, J. (2015). Relationship between participants' level of education and engagement in their completion of the Understanding Dementia Massive Open Online Course. BMC Medical Education, 15(1), 60.
15. Hammersley, M. (2013). What's wrong with ethnography?. London: Routledge.
16. Holland, F. (2016). Choosing not to reconstruct- post mastectomy: Exploring younger women's experiences. Nursing Network News, Breast Cancer Care UK.
17. Holland, F., Archer, S., & Montague, J. (2016). Younger women's experiences of deciding against delayed breast reconstruction post-mastectomy following breast cancer: An interpretative phenomenological analysis. Journal of Health Psychology, 21(8), 1688-1699.
18. Holland, F. G., Peterson, K., & Archer, S. (2018). Thresholds of size: An interpretative phenomenological analysis of childhood messages around food, body, health and weight. Critical Dietetics, 4(1), 25-35.
19. Huang, J., Kornfield, R., Szczypka, G., & Emery, S.L. (2014) A cross-sectional examination of marketing of electronic cigarettes on Twitter. Tobacco Control, 23(3), iii26-30.
20. Ignatow, G., & Mihalcea, R. (2017). Text Mining: A guidebook for the social sciences. Thousand Oaks: Sage.
21. Ingram, D. (2018, March 20) Factbox: Who is Cambridge Analytica and what did it do? Reuters, Technology News [blog post]. Retrieved from https://www.reuters.com/article/us-facebook-cambridge-analytica-factbox/factbox-who-is-cambridge-analytica-and-what-did-it-do-idUSKBN1GW07F

22. Kirk, J., & Miller, M. (1986). *Reliability and validity in qualitative research*. London: Sage

23. Knight, S., & Price, H. (2016, September 14). *Understanding the who, what and how of online learning*. Jisc [Blog post]. Retrieved from https://www.jisc.ac.uk/blog/understanding-the-who-what-and-how-of-online-learning-14-sep-2016

24. Koch, T., & Harrington, A. (1998). Reconceptualizing rigour: the case for reflexivity. *Journal of advanced nursing, 28(4)*, 882-890.

25. Kozinets, R. V. (2002). The field behind the screen: Using netnography for marketing research in online communities. *Journal of Marketing Research, 39(1)*, 61-72.

26. Kozinets, R. V. (2009). *Netnography: doing ethnographic research online*. Sage.

27. Mays, N., & Pope, C. (2000). Assessing quality in qualitative research. *Bmj, 320(7226)*, 50-52.

28. McInerney, F., Doherty, K., Bindoff, A., Robinson, A., & Vickers, J. (2018). How is palliative care understood in the context of dementia? Results from a massive open online course. *Palliative Medicine, 32(3)*, 594-602.

29. Oscar, N., Fox, P. A., Croucher, R., Wernick, R., Keune, J., & Hooker, K. (2017). Machine learning, sentiment analysis, and tweets: an examination of Alzheimer's disease stigma on Twitter. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 72(5)*, 742-751.

30. Pickering, J. D., & Swinnerton, B. J. (2017). An anatomy massive open online course as a continuing professional development tool for healthcare professionals. *Medical Science Educator, 27(2)*, 243-252.

31. Pink S., Horst, H., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2016). *Digital Ethnography*. Sage.

32. Ratcliffe, S. (2016). *Roy Amara 1925–2007, American futurologist. Oxford Essential Quotations* (4th ed.). Oxford University Press.

33. Rawlings, D., Tieman, J. J., Sanderson, C., Parker, D., & Miller-Lewis, L. (2017). Never say die: death euphemisms, misunderstandings and their implications for practice. *International Journal of Palliative Nursing, 23(7)*, 324-330.

34. Reavley, N. J., & Pilkington, P. D. (2014). Use of Twitter to monitor attitudes toward depression and schizophrenia: an exploratory study. *PeerJ, 2*, e647.

35. Reich, J., Ruiperez-Valiente, J.A. (2019) The MOOC pivot. *Science, 363(6423)*, 130-131

36. Robertshaw, D., & Cross, A. (2019a). Experiences of integrated care for dementia from family and carer perspectives: A framework analysis of massive open online course discussion board posts. *Dementia, 18(4)*, 1492-1506.

37. Robertshaw, D., & Cross, A. (2019b). Roles and responsibilities in Integrated Care for dementia. *Journal of Integrated Care, 27(2)*, 131-140.

38. Robertshaw, D., & Kotera, Y. (2019) Changing Attitudes with a MOOC on Dementia. *European Journal of Open, Distance and E-learning, 22(2)*, 27-39. Retrieved from https://www.eurodl.org/materials/contrib/2019/Robertshaw_Kotera.pdf

39. Robillard, J. M., Johnson, T. W., Hennessey, C., Beattie, B. L., & Illes, J. (2013). Aging 2.0: health information about dementia on Twitter. *PLoS One, 8(7)*, e69861.

40. Rodriguez, C. O. (2012). MOOCs and the AI-Stanford Like Courses: Two Successful and Distinct Course Formats for Massive Open Online Courses. *European Journal of Open, Distance and E-Learning, 2012(II)*. Retrieved from https://www.eurodl.org/materials/contrib/2012/Rodriguez.pdf

41. Rosewell, J., & Jansen, D. (2014). The OpenupEd quality label: benchmarks for MOOCs. *INNOQUAL: The International Journal for Innovation and Quality in Learning, 2(3)*, 88-100.

42. de Ruddere, L., Goubert, L., Prkachin, K. M., Stevens, M. A. L., van Ryckeghem, D. M. L., & Crombez, G. (2011). When you dislike patients, pain is taken less seriously. *PAIN®, 152(10)*, 2342-2347.

43. Scanfeld, D., Scanfeld, V., & Larson, E. L. (2010). Dissemination of health information through social networks: Twitter and antibiotics. *American Journal of Infection Control, 38(3)*, 182-188.

44. Shah, D. (2018 December 11). *By the numbers: MOOCs in 2018. Class Central MOOCReport* [Blog post]. Retrieved from https://www.class-central.com/report/mooc-stats-2018/

45. Siemens, G. (2005). Connectivism: A Learning Theory for the Digital Age. *International Journal of Instructional Technology and Distance Learning.*

46. Sinnenberg, L., DiSilvestro, C. L., Mancheno, C., Dailey, K., Tufts, C., Buttenheim, A. M., ... & Asch, D. A. (2016). Twitter as a potential data source for cardiovascular disease research. *JAMA cardiology, 1(9)*, 1032-1036.

47. Sudweeks, F., & Rafaeli, S. (1996). How do you get a hundred strangers to agree?. *Computer Networking and Scholarly Communication in the Twenty-first Century*, 115-136. New York: State University of New York.

48. Sveningsson, M. (2003). Ethics in internet ethnography. *Internal Journal of Global Health Information Management, 11(3)*, 45-61.

49. Tattersall, A., & Grant, M. J. (2016). Big data–what is it and why it matters. *Health Information & Libraries Journal, 33(2)*, 89-91.

50. Tribett, E. L., Ramchandran, K., Fronk, J., Passaglia, J., Bugos, K., Kogon, M., ... & Lyo, G. (2016). Palliative Care Always as a massive open online course (MOOC) to build primary palliative care in a global audience. *J Clin Oncol, 34(29)*, 123.

51. Twitter (2018a) *Privacy Policy*. Retrieved from https://twitter.com/en/privacy

52. Twitter (2018b) *Developer Agreement and Policy*. Retrieved from https://developer.twitter.com/en/developer-terms/agreement-and-policy.html