# On Algorithms, 'Big Data' and the Future of Psychometrics

Kenneth D. Royal and Melanie Lybarger



The topic of automation replacing human jobs has been receiving a great deal of media attention in recent months. In January, the McKinsey Global Institute (Manyika et al., 2017) published a report stating 51% of job tasks (not jobs) could be automated with current technologies. The topic of 'big data' and algorithms was also briefly discussed on the Rasch listserv last year and offered a great deal of food-for-thought regarding the future of psychometrics in particular. Several individuals noted a number of automated scoring procedures are being developed and fine-tuned, and each offer a great deal of promise. Multiple commenters noted the potential benefits of machine scoring using sophisticated algorithms, such as power, precision, and reliability. Some comments even predicted humans will become mostly obsolete in the future of psychometrics. Certainly, there is much to get excited about when thinking about the possibilities. However, there remain some issues that should encourage us to proceed with extreme caution.

**The Good**

For many years now algorithms have played a significant role in our everyday lives. For example, if you visit an online retailer's website and click to view a product, you will likely be presented a number of recommendations for related products based on your presumed interests. In fact, years ago Amazon employed a number of individuals whose job was to critique books and provide recommendations to customers. Upon developing an algorithm that analyzed data about what customers had purchased, sales increased dramatically. Although some humans were (unfortunately) replaced with computers, the 'good' was that sales skyrocketed for both the immediate and foreseeable long-term future and the company was able to employ many more people. Similarly, many dating websites now use information about their subscribers to predict matches that are likely to be compatible. In some respects, this alleviates the need for friends and acquaintances to make what are often times awkward introductions between two parties, and feel guilty if the recommendation turns out to be a bad one. The 'good', in this case, is the ability to relieve people that have to maintain relationships with each party of the uncomfortable responsibility of playing matchmaker.

While the aforementioned algorithms are generally innocuous, there are a number of examples that futurists predict will change most everything about our lives. For example, in recent years Google's self-driving cars have gained considerable attention. Futurists imagine a world in which computerized cars will completely replace the need for humans to know how to drive. These cars will be better drivers than humans - they will have better reflexes, enjoy greater awareness of other vehicles, and will operate distraction-free (Marcus, 2012). Further, these cars will be able to drive closer together, at faster speeds, and will even be able to drop you off at work while they park themselves. Certainly, there is much to look forward to when things go as planned, but there is much to fear when things do not.

## The Bad

Some examples of algorithmic failures are easy to measure in terms of costs. In 2010, the 'flash crash' occurred when an algorithmic failure from a firm in Kansas who ordered a single mass sell and triggered a series of events that led the Dow Jones Industrial Average into a tailspin. Within minutes, nearly $9 trillion in shareholder value was lost (Baumann, 2013). Although the stocks later rebounded that day, it was not without enormous anxiety, fear and confusion.

Another example involving economics also incorporates psychosocial elements. Several years ago, individuals (from numerous countries) won lawsuits against Google when the autocomplete feature linked libelous and unflattering information to them when their names were entered into the Google search engine. Lawyers representing Google stated "We believe that Google should not be held liable for terms that appear in autocomplete as these are predicted by computer algorithms based on searches from previous users, not by Google itself." (Solomon, 2011). Courts, however, sided with the plaintiffs and required Google to manually change the search suggestions.

Another example involves measures that are more abstract, and often undetectable for long periods of time. Consider 'aggregator' websites that collect content from other sources and reproduces it for further proliferation. News media sites are some of the most common examples of aggregators. The problem is media organizations have long been criticized with allegations of bias. Cass Sunstein, Director of the Harvard Law School's program on Behavioral Economics and Public Policy, has long discussed the problems of 'echo chambers', a phenomenon that occurs when people consume only the information that reinforces their views (2009). This typically results in extreme views, and when like-minded people get together, they tend to exhibit extreme behaviors. The present political landscapes in the United States (e.g., democrats vs. republicans) and Great Britain (e.g., "Brexit" - Britain leaving the European Union) highlight some of the consequences that result from echo chambers. Although algorithms may not be directly responsible for divisive political views throughout the U.S. (and beyond), their mass proliferation of biased information and perspectives certainly contributes to group polarization that may ultimately leave members of a society at odds with one another. Some might argue these costs are among the most significant of all.

## The Scary

Gary Marcus, a professor of cognitive science at NYU, has published a number of pieces in *The New Yorker* discussing what the future may potentially hold if (and when) computers and robots reign supreme. In a 2012 article he presents the following scenario:

*Your car is speeding along a bridge at fifty miles per hour when an errant school bus carrying forty innocent children crosses its path. Should your car swerve, possibly risking the life of its owner (you), in order to save the children, or keep going, putting all forty kids at risk? If the decision must be made in milliseconds, the computer will have to make the call.*

Marcus' example underscores a very serious problem regarding algorithms and computer judgments. That is, when we outsource our control we are also outsourcing our moral and ethical judgment.

Let us consider another example. The Impermium corporation, which was acquired by Google in 2014, was essentially an anti-spam company whose software purported to automatically "identify not only spam and malicious links, but all kinds of harmful content—such as violence, racism, flagrant profanity, and hate speech—and allows site owners to act on it in real-time, before it reaches readers." As Marcus (2015) points out, how does one "translate the concept of harm into the language of zeroes and ones?" Even if a technical operation was possible to do this, there remains the problem that morality and ethics is hardly a universally agreed upon set of ideals. Morality and ethics are, at best, a work-in-progress for humans, as cultural differences and a host of contextual circumstances presents an incredibly complex array of confounding variables. These types of programming decisions could have an enormous impact on the world. For example, algorithms that censor free speech in democratic countries could spark civil unrest among people already suspicious of their government; individuals flagged to be in violation of an offense could have his/her reputation irreparably damaged, be terminated by an employer, and/or charged with a crime(s). When we defer to computers and algorithms to make our decisions for us, we are entrusting that they have all the 'right' answers. This is a very scary proposition given the answers fed to machines come from data, which are often messy, out-of-date, subjective, and lacking in context.

An additional concern involves the potential to program evil into code. While it is certainly possible that someone could program evil as part of an intentional, malicious act (e.g., terrorism), we are referring to evil in the sense of thoughtless actions that affect others. Melissa Orlie (1997), expanding on the idea of "ethical trespassing" as originally introduced by political theorist Hannah Arendt, discusses the notion of 'ordinary evil'. Orlie argues that despite our best intentions, humans inevitably trespass on others by failing to predict every possible way in which our decisions might impact others. Thoughtless actions and unintended consequences must, therefore, be measured, included, and accounted for in our calculations and predictions. That said, the ability to do this perfectly in most contexts can never be achieved, so it would seem each day would present a new potential to open Pandora's Box.

## Extensions to Psychometrics

Some believe the 'big data' movement and advances in techniques designed to handle big data will, for the most part, make psychometricians obsolete. No one knows for sure what the future holds, but at present that seems to be a somewhat unlikely proposition. First, members of the psychometric community are notorious for being incredibly tedious with respect to not only the accuracy of information, but also the inferences made and the way in which results are used. Further, it is apparent that the greatest lessons learned from previous algorithmic failures pertains to the unintended consequences, albeit economically, socially, culturally, politically, and legally that may result (e.g., glitches that result in stock market plunges, legal liability for mistakes, increased divisions in political attitudes, etc.). Competing validity conceptualizations aside, earnest efforts to minimize unintended consequences is something most psychometricians take very seriously and already do. If anything, it seems a future in which algorithms are used exclusively could only be complemented by psychometricians who perform algorithmic audits (Morozov, 2013) and think meticulously about identifying various 'ordinary evils'. Perhaps instead of debating whether robots are becoming more human or if humans are becoming more robotic, we would be better off simply appreciating and leveraging the strengths of both?

## References

Baumann, N. (2013). Too fast to fail: How high-speed trading fuels Wall Street disasters. *Mother Jones*. Available at: http://www.motherjones.com/politics/2013/02/high-frequency-trading-danger-risk-wall-street

Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., & Dewhurst, M. (2017). A future that works: Automation, employment, and productivity. The McKinsey Global Institute. Available at: http://www.mckinsey.com/global-themes/digital-disruption/harnessing-automation-for-a-future-that-works

Marcus, G. (2012). Moral machines. *The New Yorker*. Available at: http://www.newyorker.com/news/news-desk/moral-machines

Marcus, G. (2015). Teaching robots to be moral. *The New Yorker*. Available at: http://www.newyorker.com/tech/elements/t robots-to-be-moral

Morozov, E. *To Save Everything, Click Here: The Folly of Technological Solutionism* (2013). PublicAffairs Publishing, New York, NY.

Orlie, M. (1997). *Living ethically, acting politically*. Cornell University Press, Ithaca, NY.

Solomon, K. (2011). Google loses autocomplete lawsuit. *Techradar*. Available at: http://www.techradar.com/news/internet/g loses-autocomplete-lawsuit-941498

Sunstein, C. R. (2009). *Republic.com 2.0*. Princeton University Press, Princeton, NJ.