Tal Yarkoni

Department of Psychology and Neuroscience

University of Colorado Boulder

UCB 345

Boulder, CO 80309

Tel: (303) 492-4299

E-mail: tal.yarkoni@colorado.edu

Word count: 2,994 words

Psychoinformatics: New horizons at the interface of the psychological and computing sciences

Tal Yarkoni[1]

[1] Institute of Cognitive Science and Department of Psychology and Neuroscience, University of Colorado Boulder

Abstract

Psychologists live in an increasingly data-rich world, and the ability to make continued progress in understanding the mind and brain depends on finding new ways to organize and synthesize an ever-expanding body of knowledge. Here I review current research in psychoinformatics—an emerging discipline that uses tools and techniques from the computer and information sciences to improve the acquisition, organization, and synthesis of psychological data. I focus on several areas where the application of informatics approaches has already paid large dividends, including novel data collection approaches; adaptation of computational techniques and insights; aggregation and organization of psychological data; large-scale data mining and synthesis; and improving research and publication practices. I argue that in the coming years, informatics approaches are likely to play the same instrumental role in shaping psychological research that they have already played in other fields such as genetics and neuroscience.

Keywords: Informatics, Methods, Data mining, Information science

Introduction

Scientific progress increasingly depends our ability to harness and apply tools and techniques from the computer and information sciences. Many major scientific advances—for example, the Human Genome Project and the Large Hadron Collider—depend critically on technical advances in the large-scale acquisition, management, and synthesis of scientific data. This application of methods from the computer and information sciences to other fields of science isn't just a happy accident; it's also a field in its own right—one commonly referred to as *informatics*. Prefix that term with a Greek root or two and you get other terms like *bioinformatics*, *neuroinformatics*, and *ecoinformatics*—all well-established fields responsible for many of the most exciting recent discoveries in their parent disciplines.

Curiously, following the same convention also gives us a field called *psychoinformatics*—which, if you believe Google Scholar, barely exists at all. A search for the term returns only 18 hits as of this writing, compared to > 1 million, 18,000, and 3,000 for bioinformatics, neuroinformatics, and ecoinformatics, respectively. The discrepancy is surprising, because labels aside, it's clear that psychological scientists are already harnessing informatics methods in powerful and creative ways, often reshaping the very way we collect, organize, and synthesize our data. In the present article, I review some of these recent developments. To emphasize the utility of informatics approaches at all stages of the research process, the paper is structured in roughly the same way that a typical study might proceed—beginning with data collection, proceeding to analysis and organization of results, and culminating with communication and evaluation of findings. The overarching argument is that psychoinformatics is, for all intents and purposes, already a vibrant field making important contributions to psychological research, and that it is in psychological scientists' best interest to formally recognize it as such so as to encourage its further development.

Redefining the psychology lab

The process of collecting data from human participants used to take place predominantly in dedicated, physical lab spaces, but improvements in technology have recently expanded the boundaries of the psychology lab in several ways. The emergence of the Web has enabled psychological scientists to recruit internet-based samples, which are typically more diverse, larger, and cheaper to acquire than conventional samples (Gosling, Vazire, Srivastava, & John, 2000). While there are many things one can't do online (e.g., measure physiological responses), the availability of an effectively limitless userbase has fundamentally altered the data collection landscape in many domains. Easy-to-use web applications like SurveyMonkey (http://surveymonkey.com) and Qualtrics (http://qualtrics.com) allow anyone to create and administer sophisticated surveys, while researchers with programming experience can use languages like JavaScript to create online analogs of most offline experiments. Participant recruitment is greatly facilitated by services like Amazon's Mechanical Turk (MTurk; http://mturk.com)—a marketplace that allows researchers to obtain relatively high-quality data from thousands of participants at very low cost (Buhrmester, Kwang, & Gosling, 2011).

In parallel to the growth of the Web, mobile consumption of data services via 'smart' devices has exploded. By the end of 2011, there were more than 1.2 billion mobile broadband subscriptions worldwide (International Telecommunications Union, 2011); in many developed countries, over half of mobile users now own smartphones. We're told there's an app for everything, and psychological science is no exception. In the past few years, researchers have used smartphone applications to study a range of psychological phenomena. For example, Killingsworth and Gilbert (2010) used an iPhone experience sampling application to quantify the proportion of time people spend daydreaming and demonstrate that people are less happy when daydreaming than engaging in other activities. Dufau and colleagues developed a lexical decision task application with seven language-specific versions, enabling the authors to nearly effortlessly collect data from over 4,000 participants in just four months (Dufau et al., 2011). Mehl and colleagues have developed an innovative portable recording device—and now also an iPhone app—called the EAR (Electronically Activated Recorder) that periodically records audio snippets of people's daily lives (Mehl, Pennebaker, Crow, Dabbs, & Price, 2001). Data from the EAR has taught us many interesting things about what people do when they're not in the lab—for example, that contrary to popular opinion, women do not talk more than men (Mehl, Vazire, Ramírez-Esparza, Slatcher, & Pennebaker, 2007), and that close others can predict many daily behaviors better than the self can (Vazire & Mehl, 2008). In the long term, experience sampling via mobile applications seems likely to assume a central role in psychological data collection (Miller, 2012).

In addition to facilitating new data collection, technology also provides researchers an unparalleled window into people's mental lives via existing datasets. In the simple act of living, many of us generate a continuous stream of information: We text our friends, track our locations with GPS, upload pictures of our activities, and stream music and movies through the air. Much of this data now persists indefinitely in our online accounts, and can often be programmatically accessed via the web, providing social scientists with the ability to analyze real-world behavior on an unprecedented scale. For example, Golder and Macy (2011) analyzed millions of public tweets to quantify changes in mood in relation to time of day and day of the week. Gosling and colleagues used Facebook profiles to demonstrate that participants' Extraversion levels predict their number of friends very strongly (Gosling, Augustine, Vazire, Holtzman, & Gaddis, 2011). Yarkoni (2010a) used public data from Google's Blogger service to conduct a large-scale investigation of the role of personality in shaping word use in nearly 700 online participants' blogs, comprising nearly 80 million words (Figure 1A). Ten or twenty years ago, such investigations would have been logistically impossible; today, they can be accomplished in a matter of days or weeks by individual researchers. And such efforts to harness ‘Big Data’ undoubtedly only scratch the surface. Some of the most exciting psychological applications of technology are likely to involve new kinds of data that lack conventional analogs—e.g., GPS-based location data, readings from Bluetooth-enabled physiological sensors, and whole-genome scans obtained through personal genomics services (e.g., 23andMe.com).

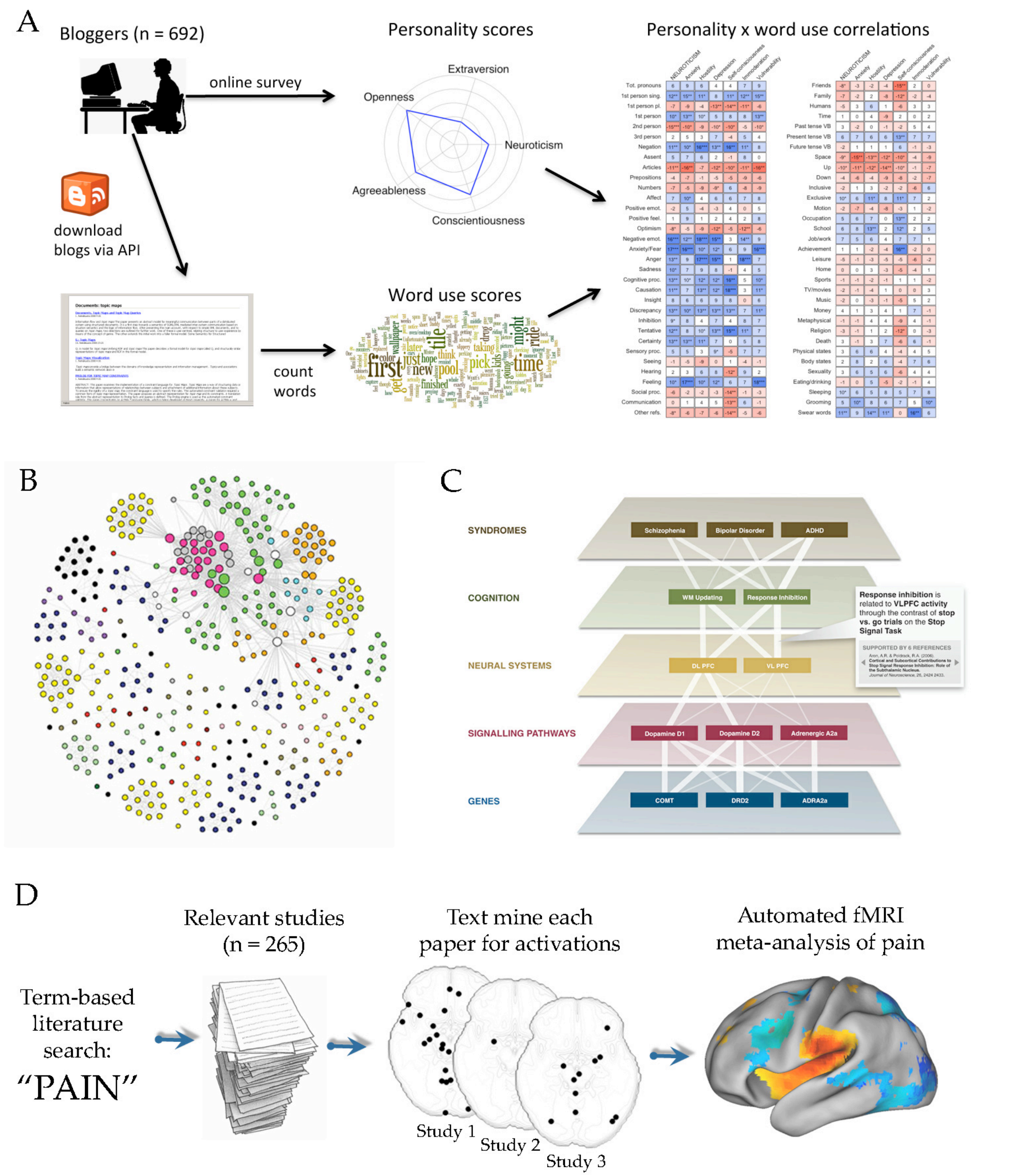

Figure 1. Examples of recent psychoinformatic applications. (A) In Yarkoni (2010a), bloggers filled out an online personality questionnaire and provided their blog address. Google's Blogger API was used to automatically download all blog contents. Word count-based measures of language use were then extracted from each participant's blog, enabling large-scale investigation of the role of personality in shaping word use. (B) Network analyses reveal a small-world structure to DSM-IV symptoms (from Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011). Each node represents a symptom, with edges connecting nodes that co-occur in at least one disorder. (C) Schematic illustration of a formal ontology that enables mapping of relations between psychological and biological constructs at different levels (from Poldrack et al., 2011). (D) Schematized illustration of approach used to automatically meta-analyze large subsets of the fMRI literature in (Yarkoni et al., 2011). Articles that use the target term (e.g., 'pain') at high frequency are selected for inclusion; a custom text parser is then used to automatically extract all reported brain activations from all articles, and the results are subjected to a standard fMRI meta-analysis (Wager, Lindquist, & Kaplan, 2007).

A growing role for computational analysis

Computational methods play an increasingly important role in the social sciences (Lazer et al., 2009), and have much to offer psychological scientists as well. The widespread availability of high-end computing resources now facilitates many once-impractical applications—for example, resampling-based statistical analyses that require none of the assumptions associated with many of the most widely used parametric tests (Siegel, 1957). Modern software packages allow psychological scientists to perform an array of research-related tasks—from data cleaning to visualization—quickly and elegantly. A prominent example is the free open-source statistical analysis software *R,* which features hundreds of user-contributed packages supporting everything from data reshaping to structural equation modeling to multivariate behavioral genetic analyses (R Development Core Team, 2010).

A particularly useful set of computational techniques and concepts is found in the *machine learning* literature (Alpaydin, 2010), which focuses on developing prediction and classification algorithms that can learn from experience. Machine learning approaches are already widely used in the neuroimaging literature to classify and decode mental states from brain activity (Pereira, Mitchell, & Botvinick, 2009); similar approaches could have broad applicability in other domains, particularly in clinical applications involving the prediction and classification of mental health disorders. Machine-learning approaches are particularly helpful in overcoming the widespread problem of *overfitting*—i.e., the tendency of most statistical models to capitalize on chance and thereby produce overly optimistic effect size estimates (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009).

More generally, adapting tools and techniques from the computer and information sciences can help improve the measurement and modeling of psychological processes in a broad range of ways. For example, Borsboom and colleagues used standard network analysis techniques to show that the symptoms of most DSM-IV disorders exhibit a small-world structure (a common network structure in which the vast majority of nodes are connected to each other through a small proportion of highly connected 'hubs'; Fig. 1B), in contrast to the traditional conceptualization in terms of discrete entities (Borsboom et al., 2011). Yarkoni and colleagues introduced a new measure of orthographic similarity (i.e., an index of how visually similar each word is to other words) based on a standard computer science metric called the Levenshtein distance, and showed that the new measure predicts visual word recognition performance substantially better than previous measures (Yarkoni, Balota, & Yap, 2008). And Yarkoni (2010b) developed a novel method for automatically abbreviating questionnaire measures (e.g., personality inventories) by using a genetic algorithm—a standard computer science technique that draws on evolutionary principles to 'evolve' progressively good solutions to complex problems. The key point is that such applications need not require new technical innovation; their utility lies primarily in the cross-disciplinary adaptation of standard computational techniques to open problems in psychological science.

Aggregating and organizing the data

A considerable proportion of informatics research in other disciplines focuses on assembling very large databases and providing structured access to their contents. There are promising signs of a growing focus on large-scale data organization and aggregation within psychological science as well. First, researchers are devoting increasing resources to assembling *megastudies*—experimental datasets that sample from an unusually large number of participants and/or stimuli (Balota, Yap, Hutchison, & Cortese, 2012). One prominent example is the English Lexicon Project (ELP) developed by Balota and colleagues, which contains word naming and lexical decision data for nearly 40,000 English words (Balota et al., 2007). Another example is YourMorals.Org, a data collection platform developed by Iyer and colleagues that focuses on moral psychology studies, with over 200,000 participants to date. The scope of such datasets has enabled researchers to answer questions that would be difficult or impossible to address using conventional lab-based approaches (e.g., Graham et al., 2011; New, Ferrand, Pallier, & Brysbaert, 2006).

Second, many psychology datasets are now freely available from online databases. Though databases targeted at psychological scientists remain scarce, hundreds of psychology datasets can be downloaded from social science-wide databases such as the IQSS Dataverse Network (dvn.iq.harvard.edu) or domain-specific databases such as the Open fMRI project (openfmri.org). Such resources allow researchers to verify, extend and synthesize other researchers' results with greater efficiency.

Ultimately, to benefit maximally from publicly accessible datasets, psychological scientists will need to develop comprehensive ontologies of psychological constructs that allow data to be queried and retrieved in structured ways (Fig. 1C). Incipient efforts are already underway; for example, the Cognitive Atlas Project (http://cognitiveatlas.org) seeks to build a collaborative knowledge base of cognitive concepts and tasks, with the aim of facilitating "intelligent aggregation of research findings" (Poldrack et al., 2011). Although a consensus ontology encompassing all of psychological science remains a distant prospect, even rudimentary ontologies should greatly facilitate psychologists' ability to organize, share, and access the fruits of their collective labor.

Large-scale data mining and synthesis

No matter how data are obtained, one must ultimately synthesize them into a meaningful form. Here, again, psychological scientists can profitably adapt approaches pioneered in other fields. For instance, much of bioinformatics has focused on developing tools for large-scale genomic data mining. Although the potential for truly comprehensive data mining applications is somewhat constrained by the relatively unstructured nature of psychological data, there are nevertheless plenty of opportunities for more modest applications. For example, my colleagues and I recently developed a platform called Neurosynth (http://neurosynth.org; Fig. 1D; Yarkoni et al., 2011) that uses data reported in thousands of published neuroimaging articles to automatically generate fMRI meta-analyses and "decode" mental states from patterns of brain activity.

Perhaps the greatest promise of a large-scale data mining approach is its potential to drive truly exploratory science. Psychological science is often presented as a confirmatory, hypothesis-driven enterprise (arguably to an exaggerated extent—see Bones, 2012), but we can already glimpse a future in which exploratory data analysis plays a complementary role in generating entirely novel hypotheses. The combination of enormous datasets, richer experience-sampling techniques, and machine learning algorithms may soon enable us to identify novel associations in a purely data-driven way—much as genome-wide association studies (GWAS) have revolutionized complex disease genetics. One early example is brainSCANr (http://brainscanr.com), a text mining engine designed to uncover potentially unpredicted relationships between neurobiological constructs (Voytek & Voytek, 2012). Such exploration platforms will present novel challenges (e.g., the need to correct for thousands of simultaneous statistical comparisons) and raise new ethical considerations (e.g., preserving the privacy of social media users who haven't explicitly consented to have their data used), but have the potential to identify many effects that may simply be too counterintuitive to discover via traditional routes—for example, the initially puzzling finding that redheads have differential pain tolerance, which subsequent investigation suggests reflects variation in the melanocortin-1 receptor gene (Mogil et al., 2005).

Evaluating and communicating results

Finally, an increased focus on harnessing technology could substantially improve procedural aspects of psychological research. Psychological science is an imperfect enterprise: replication is under-emphasized, analysis and reporting standards can be lax (Simmons, Nelson, & Simonsohn, 2011), and conventional peer review is unreliable (Bornmann, Mutz, & Daniel, 2010). There's a tremendous incentive to develop tools and platforms that can help address such problems. One important direction is to adapt collaborative filtering algorithms widely used on commercial and social website to improve the post-publication review of articles (Nosek & Bar-Anan, in press; Yarkoni, in press). Another is to develop databases that facilitate tracking of null results and other unpublished findings—such as the recently introduced PsychFileDrawer site (http://psychfiledrawer.org). Further down the road, one can envision many other applications, such as automated quality-control algorithms that detect patterns suggestive of publication bias (e.g., too many p values just below $p < .05$ in a researcher's published output), or personalized recommendation systems that can identify articles likely to be of interest to individual researchers.

Improvements in technology are also driving rapid changes in the way psychological scientists communicate their results to one another as well as to the broader public. In some cases, the push to leverage technology has come from the top down, as in APS's recent Wikipedia initiative calling on APS members to help improve the accuracy of psychology entries in Wikipedia (www.psychologicalscience.org/apswi). More commonly, however, such changes are occurring organically, as researchers spontaneously discover the benefits of new media technology. Already, hundreds of academic psychologists (including the present author) maintain blogs or Twitter streams, and the past few years have seen a proliferation of social networks targeted at professional scientists (e.g., ResearchGATE and Mendeley). These tools provide powerful new ways to share and evaluate new findings, request feedback, and communicate results to the general public, and they are likely to see increasing adoption within the scientific community as time goes on.

Conclusion

In the title of a provocative recent article, Greenwald asserted that "There is nothing so theoretical as a good method" (Greenwald, 2012)—alluding to the fact that despite the preeminence of theory and theoretical controversy in experimental psychology, a disproportionate number of lasting scientific contributions are actually methodological rather than theoretical in nature. Given the increasingly central role of technology in science, this trend seems likely to strengthen rather than abate, with innovations in the computer and information sciences continuing to translate into exciting new discoveries in the psychological sciences.

Fortunately, psychological scientists are already in an excellent position to take advantage of such advances. Although the term *psychoinformatics* may be underutilized, a considerable amount of work reviewed above already falls under this rubric. What is needed now is further community and institutional support to help solidify psychoinformatics as a full-fledged discipline. In the short term, individual researchers can contribute to this goal by, for example, taking advantage of novel data collection techniques, switching from proprietary software packages like SPSS to open-source environments like *R*, and investing the time and effort to learn new computing skills. Psychologists who already have strong technical backgrounds can pursue informatics-related funding opportunities (e.g., in the US, the NSF's Office of Cyberinfrastructure routinely announces relevant programs) or explore the growing number of research opportunities in industry—where social media companies like Facebook and Twitter often have access to datasets and resources that academic psychologists can only dream of.

In the longer term, the success of psychoinformatics as an independent discipline is likely to hinge on the development of graduate training programs designed to provide budding psychoinformaticians with a strong grounding in both substantive psychological science and relevant areas of the information and computer sciences. In addition to standard training in statistics and experimental design, such training programs would require and offer coursework in software development, online data collection, machine learning, and large-scale data analytics. Graduates of such programs would be in a unique position to acquire rich new datasets on an unprecedented scale; to efficiently explore and synthesize huge amounts of information while minimizing the opportunity for human bias; and to come up with new ways of reporting and disseminating psychological findings to other scientists and to the public at large. If the trajectory of fields like biology and neuroscience is any guide, such developments will pay enormous dividends. We should embrace psychoinformatics as a full-fledged discipline and work to ensure that psychological science remains a vibrant, forward-looking field, ready to benefit from technological innovations as soon as they emerge.

References

Alpaydin, E. (2010). *Introduction to Machine Learning*. (T. Dietterich, Ed.) (p. 584). MIT Press.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., et al. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445-459.

Balota, D. A., Yap, M. J., Hutchison, K. A., & Cortese, M. J. (2012). Megastudies: What do millions (or so) of trials tell us about lexical processing? In J. S. Adelman (Ed.), *Visual word recognition volume 1: Models and methods, orthography and phonology*. Psychology Press.

Bones, A. K. (2012). We Knew the Future All Along: Scientific Hypothesizing is Much More Accurate Than Other Forms of Precognition--A Satire in One Part. *Perspectives on Psychological Science*, *7*(3), 307-309.

Bornmann, L., Mutz, R., & Daniel, H.-D. (2010). A reliability-generalization study of journal peer reviews: a multilevel meta-analysis of inter-rater reliability and its determinants. *PloS one*, *5*(12), e14331.

Borsboom, D., Cramer, A. O. J., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The Small World of Psychopathology. *PLoS ONE*, *6*(11), e27407.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). *Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? Perspectives on Psychological Science* (Vol. 6, pp. 3-5).

Dufau, S., Duñabeitia, J. A., Moret-Tatay, C., McGonigal, A., Peeters, D., Alario, F. X., Balota, D. A., et al. (2011). Smart Phone, Smart Science: How the Use of Smartphones Can Revolutionize Research in Cognitive Science. *PLoS ONE*, *6*(9), e24974.

Golder, S. A., & Macy, M. W. (2011). Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. *Science*, *333*(6051), 1878-1881.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2000). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist*, *59*(2), 93-104.

Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N., & Gaddis, S. (2011). Manifestations of personality in online social networks: self-reported facebook-related behaviors and observable profile information. *Cyberpsychology behavior and social networking*, *14*(9), 483-488.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366-385.

Greenwald, A. G. (2012). There is nothing so theoretical as a good nethod. *Perspectives on Psychological Science*, *7*(2), 99-108.

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, *330*(6006), 932.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, *12*(5), 535-540.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-lászló, Brewer, D., Christakis, N., et al. (2009). Computational Social Science. *Science*, *323*(February), 721-723.

Mehl, M. R., Pennebaker, J. W., Crow, D. M., Dabbs, J., & Price, J. H. (2001). The Electronically Activated Recorder (EAR): a device for sampling naturalistic daily activities and conversations. *Behavior Research Methods*, *33*(4), 517-523.

Mehl, M. R., Vazire, S., Ramírez-Esparza, N., Slatcher, R. B., & Pennebaker, J. W. (2007). Are women really more talkative than men? *Science*, *317*(5834), 82.

Miller, G. F. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*.

Mogil, J. S., Ritchie, J., Smith, S. B., Strasburg, K., Kaplan, L., Wallace, M. R., Romberg, R. R., et al. (2005). Melanocortin-1 receptor gene variants affect pain and mu-opioid analgesia in mice and humans. *Journal of Medical Genetics*, *42*(7), 583-587.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project, *13*(1), 45-52.

Nosek, B. A., & Bar-Anan, Y. (in press). Scientific Utopia: I. Opening scientific communication. *Psychological Inquiry*.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage*, *45*(1), S199-S209.

Poldrack, R. A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., Parker, D. S., et al. (2011). The Cognitive Atlas: Toward a Knowledge Foundation for Cognitive Neuroscience. *Frontiers in neuroinformatics*, *5*(September), 11.

R Development Core Team. (2010). R: A Language and Environment for Statistical Computing. (Austria, Ed.)*R Foundation for Statistical Computing Vienna Austria*. R Foundation for Statistical Computing.

Siegel, S. (1957). Nonparametric Statistics. *American Statistician*, *11*(3), 13-19.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. (H. E. Nissen, H. K. Klein, & R. Hirschheim, Eds.)*Psychological Science*, *22*(11), 1359-1366.

International Telecommunications Union (2011). Key Global Telecom Indicators for the World Telecommunication Service Sector. Retrieved from http://www.itu.int/ITU-D/ict/statistics/at_glance/KeyTelecom.html

Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: the accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, *95*(5), 1202-16.

Voytek, J. B., & Voytek, B. J. (2012). Automated Cognome Construction and Semi-automated Hypothesis Generation. *Journal of Neuroscience Methods*.

Wager, T. D., Lindquist, M., & Kaplan, L. (2007). Meta-analysis of functional neuroimaging data: current and future directions. *Social Cognitive and Affective Neuroscience*, *2*(2), 150-158.

Yarkoni, T. (2010a). Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality*, *44*(3), 363-373.

Yarkoni, T. (2010b). The Abbreviation of Personality, or how to Measure 200 Personality Scales with 200 Items. *Journal of Research in Personality*, *44*(2), 180-198.

Yarkoni, T. (in press). Designing next-generation platforms for evaluating scientific output: What scientists can learn from the social web. *Frontiers in Computational Neuroscience*.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Revew*, *15*(5), 971.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Essen, V., David, C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*(8), 665-670.

Recommended Readings

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2000). (See References). A classic article empirically refuting common myths about the viability of conducting psychological research on the internet.

Lazer, David et al. (2009). (See References). An accessible review of the benefits of a computational approach to the social sciences.

Miller, G. F. (2012). (See References). An aptly-titled manifesto outlining current and future capabilities for smartphone-based psychological research.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). (See References). Introduces a highly automated text-mining framework for synthesizing and decoding the results of thousands of published functional neuroimaging articles.