

“Big Data” Psychology

Karsten Hansen

UC San Diego
The Rady School of Management

January 2014

Abstract

In every aspect of our daily lives, from the way we work, shop, communicate, or socialize, we are both consuming and creating vast amounts of information. More often than not, these daily activities create a trail of digitized data that is stored, mined, and analyzed by firms hoping to create valuable business intelligence. This talk will focus on how such large databases can provide useful insights for research in social psychology. My basic contention is that when aggregated over a large number of individuals, seemingly innocuous activities such as a web search or a casual trip to a neighborhood store can reflect aspects of our deep-rooted ideologies, values, and personality traits. The talk will focus on three separate case-studies.

The first study focuses on Habituation and Variety Seeking. An extensive literature in psychology and cognitive science documents that repetitive behaviors (and thoughts) in a stable context can take the form of automated reflexive responses to external or internal triggers, i.e. take the form of habits. A counter argument to habitual behavior is an extensive literature documenting intrinsic motivations to seek variety even in the presence of the most favorite option. We study habituation and variety seeking in one of the most mundane activities of life: Grocery shopping. We rely on a unique database that records the details of grocery shopping histories for a nationwide demographically balanced sample of over 100,000 households observed over several years. For each house hold on each shopping occasion, we observe the details of the product chosen, price paid, and indicator of any promotional activities. We utilize household transaction histories to compute measures of purchase concentrations such as Herndahl Index (HHI) and concentration ratios (CR) in a wide variety of food and non-food products. The data includes a wide array of demographics such as income, education, age and household size. Although we do not have measures on personality traits, our strategy is to create theoretically relevant proxies using external data.

The second study focuses on Aging and Decision making. In this work, we investigate the extent to which the age-related cognitive differences observed in controlled lab studies also manifest in the trivial choices of our daily lives, such as a casual trip to the neighborhood supermarket. For example, if older adults have smaller, but more meaningful social networks and spheres of interests, does this persist beyond the social and personal domains, such that older consumers also show higher brand affinity and a narrower product portfolio of grocery products? Given that the largest age related differences in lab studies are in the domain of novel and uncertain tasks, would similar differences manifest in trivial decisions such as willingness to try a new flavor of

yogurt? Similarly, risk aversion is found to increase with age, particularly when weighing potential losses. Can risk aversion and a general prevention focus in the elderly generate avoidance of generic products as opposed to national brands, particularly in riskier choice of a (chemically identical) generic pill over branded Tylenol? Our empirical strategy involves creating theoretically relevant measures to identify age-related cognitive differences as manifested in purchase behavior. We use three key measures: (1) new product trial, (2) concentration of purchases, and (3) use of branded products versus generic alternatives. Each of these measures captures a distinct psychological phenomenon. For example, the propensity to try new products capture aspects of novelty in an unfamiliar situation (e.g. trying a new brand of soda), while product concentration measures habituation or variety seeking behavior amongst existing options with which consumer has considerable experience. The use of branded versus generic alternatives is used to capture risk aversion. Industry reports and academic research in marketing suggests that generics are perceived by consumers as riskier and of lower quality than national brands.

In the final study our objective is to explain variation in life satisfaction for about 1.4 million Americans. A large literature in both economics and psychology has sought to uncover the “happiness production-function” of individuals. Most of this literature has relied on archival data of modest dimensions and is often based on simple correlations. In this study we pool multiple waves of the largest survey in the world: The Behavioral Risk Factor Surveillance System collected annually by the Centers for Disease Control and Prevention. This provides us with a sample much larger than previously used. Using this sample we replicate some of the effects found in the previous literature, but also document a range of new findings.