Examining Fans’ Emotions in the Twittersphere: A Longitudinal Text Analysis of Fans’ Tweets Yielded During Games Between Prominent NCAA Division I FBS Rivals

Kiernan Gordon, University of New England
Christina Herrick, University of New Hampshire
Ryan Boyd, University of Texas

Thursday, May 30, 2019
Poster
Abstract 2019-083

Scholarship examining different actors’ and organizations’ use of social media within a variety of sport contexts has grown significantly (Pedersen, 2014) in recent years. Twitter has been examined substantially more than any other platform within this stream of research (Abeza, O'Reilly, Seguin, & Nzindukiyimana, 2015). Recent scholarship has begun to aggregate large Twitter datasets for various forms of textual analyses (Chang, 2018; Yu & Wang, 2015). This research, however, has emphasized cross-sectional studies and has often examined results through non-sport specific frameworks. Moreover, research findings involving smaller data sets have often been limited by sample size, corresponding metadata, and analytical tools to which scholars have had access. Consequently, more diverse methods for both data collection and analysis are needed in order to develop relevant frameworks that can effectively shape future sport scholarship (Abeza et al.), especially relative to fans’ use of Twitter (Yoon, Petrick, & Backman, 2017).

The current study represents an attempt to address these gaps via a longitudinal analysis of the emotional and linguistic content of fans’ tweets within a rivalry context. We aggregated a large sample of real-time data in the current study through Twitter’s application programming interface (API) using a predetermined set of keywords for games between two prominent NCAA Division I FBS rivals—Ohio State University and University of Michigan—in three consecutive seasons. The use of Twitter’s API allows for a “natural experiment” (Yu & Wang, 2015, p. 393) whereby fans’ psychological dynamics can be harnessed as events unfold in real-time.

Linguistic Inquiry and Word Count (LIWC; Pennebaker & Francis, 1999) has been used to analyze the tweets in the current study. LIWC has been used in sport-related Twitter research previously (see Lee & Kahle, 2016) and has been extensively validated across thousands of studies (see Tausczik & Pennebaker, 2010). Since its creation, the LIWC approach to measuring psychological information from natural language samples has been widely adopted across several fields (Boyd, 2017) while being extended to research that examines social media in recent years.

The first game in this study occurred in 2016 and served as a pilot for both data collection and analysis. 820,840 tweets were collected over a 62-hour period prior to, during, and after that game. Data was collected for the second game in November 2017. Approximately 4.9 million tweets were collected during a similar time period for the second game. Data will be collected in late November 2018 for which a similar yield is expected. Preliminary data analyses most notably suggest a tremendously high presence of the first person plural pronoun ‘we’ relative to LIWC benchmarks just prior to the start of each rivalry contest, which has implications for BIRGing (Galdini, Borden, Thorne, Walker, Freeman, & Sloan, 1976). The authors anticipate that the data from the current study will contribute insights to previous social media scholarship within sport relative to language use and emotion while further refining the most effective and cost-efficient approach to the collection of large Twitter data sets and corresponding analyses available today.