Motions for Class Certification: How Experts Can Assist the Court by Evaluating Quantitative Evidence

Palmer Morrel-Samuels, MA, M.Phil, PhD., MSc in Law

University of Michigan and Employee Motivation & Performance Assessment

ABSTRACT: Expert witnesses who are research psychologists can provide critical assistance to judges evaluating class certification motions under FRCP Rule 23. This new methodology is especially important for attorneys facing or presenting plausible but contested motions for class certification. As expert witnesses, research psychologists can apply reliable well-established methods in two ways: 1) Using double-blind independent coders to classify text-based data about proposed class members so that courts and attorneys receive empirical data on commonality and the appropriate number of subclasses within that class. 2) Using statistical methods to test quantitative data, and to formulate new validated variables that quantify the proposed representative's typicality. This unique empirical evidence regarding commonality, numerosity, and typicality will help both attorneys and judges alike to formulate, or evaluate, class certification motions. The method has withstood peer-review and a Daubert challenge in the 8th Circuit. An empirical example from case law is discussed. [1]

Research psychologists currently assist courts in many types of litigation by furnishing empirical evidence based on statistical analysis and behavioral science research (e.g., Monahan and Walker, 2014[2]). For example, research psychologists serve as expert witnesses by quantifying trademark confusion, running statistical tests of complex data, and analyzing evidence of discrimination. However, they have not always been welcomed in court, with Justice Powell denigrating their methods as "numerology derived from statistical studies" (Ballew v. Georgia, 435 U.S. 223, 1978), and Justice Roberts rejecting arguments based on behavioral science as "sociological gobbledygook." (Gill v. Whitford, 585 U.S. ____, 2018). More supportively, about 20 years ago in the Reference Manual on Scientific Evidence, Justice Breyer predicted courts will increasingly rely on statistics and behavioral science (Breyer, 2000[3]). His prediction is being borne out, as shown by research at the upcoming Conference on Empirical Legal Studies: In the 306,895 discrimination lawsuits since 1964, the primary statistical tool of many research psychologists – multiple regression – was cited in roughly 700

---

[1] This paper is drawn from an earlier presentation containing similar content: Paper presented to the Conference of the American Psychology-Law Society, Portland OR, 3/15/19.

[2] Monahan, J., & Walker, L. (1994). University casebook series. Social science in law: Cases and materials (3rd ed.). Westbury, NY, US: Foundation Press.

[3] Breyer, S., Reference Manual on Scientific Evidence, Second Edition, Introduction, 2000.

decisions, and the frequency with which those decision have been affirmed is increasing at a statistically significant rate (Morrel-Samuels, 2018[4]). Accordingly, the time seems ripe for informing attorneys and psychologists about a promising new legal application of behavioral science: assisting courts by providing unique scientific evidence to support or refute motions seeking class certification under FRCP Rule 23. This paper follows the model established by my work on the first (and currently only) litigation using this method: Robinson v. Des Moines Public Schools, LACL 136651 (2018).

§1) Quantifying Communality by Having Double-Blind Independent Coders Classify Text-based Records

This process has five steps: 1) The expert witness (EW) and a research assistant (RA) identify and isolate one or more anonymized data fields containing text about members of the proposed class; this is the corpus of text that will be categorized according to recognized principles in research psychology (e.g. Remler & Van Ryzin, 2011[5]); to expedite coding, the corpus is broken into discrete "units of analysis" where each sentence (or phrase) is demarked by punctuation (e.g., Singleton & Straits, 2010[6]). 2) The EW and RA independently review the corpus to create a list of coding categories that will allow each sentence to be classified according to categories that are straightforward and explicitly described by a brief set of coding instructions (e.g., Groves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2009[7]); coding categories are exhaustive and mutually exclusive, preferably with a one-word label marking each category (e.g., Fowler, 2009[8]). 3) Working independently the EW and RA classify each unit in the corpus using the agreed-upon categories as specified by the coding instructions. 4) The EW and RA reconvene, share their work, and collaborate as peers to reach consensus on any unit where coding differences occurred. If the number of coding differences was substantial, i.e., typically greater than 20%, then the EW may decide to revise the coding instructions until the reliability coefficient, e.g., Kappa, indicates sufficient reliability to meet publication standards in an APA journal. (McHugh, 2012[9]). 5). Once the corpus has been categorized using this double-blind independent coding process, the EW compiles a report citing the empirical evidence that does or does not support a claim of communality. The outcome of this coding process will also

---

[4] 27)   Morrel-Samuels, P. (2018) Statistical Analysis in Employment Discrimination: Trends and Implications. Presented to the Conference on Empirical Legal Studies, Society of Empirical Law Studies, Ann Arbor, 11/9/18.
[5] Remler, D. K., & Van Ryzin, G. G. (2011). Research methods in practice: Strategies for description and causation. Thousand Oaks, CA, US: Sage Publications, Inc.
[6] Singleton, R. A., & Straits, B. C. (2010). Approaches to social research. (5th Edition). New York: Oxford University Press.
[7] Groves, Robert M., Floyd J. Fowler Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer and Roger Tourangeau, 2009, Survey Methodology. Second Edition, New Jersey: John Wiley & Sons Inc.
[8] Floyd J. Fowler Jr. Survey Research Methods, Fifth Edition, Sage, 2009.
[9] McHugh M. L. (2012). Interrater reliability: the kappa statistic. Biochemia medica, 22(3), 276–282.

show if any subclasses of potential plaintiffs exist, and if so, how many subclasses should be established within the proposed class.

§2) Using Statistical Tests to Evaluate Typicality

This process has four steps: 1) The EW identifies, isolates, and anonymizes data fields containing quantitative data on the representative and the plaintiffs of the proposed class. 2) To supplement existing numerical data, the EW may construct new quantitative variables from nominal data, such as job description, degrees earned, etc. (e.g., Rosenthal & Rosnow, 2007[10]). 3) The EW then tests the reliability and validity of any constructed variables to determine if they are replicable and accurate. 4) By using conventional statistical tests supported by both case law (Hazelwood School District v. US, 433 U.S. 299, 1977) and research in psychology (Cohen, Cohen, Aiken and West, 2002[11]), the EW then determines the number of standard deviations between the proposed representative and the average of the remaining plaintiffs. To the extent that the proposed representative is found to be different from the remaining plaintiffs by a statistically significant margin, claims of his or her typicality will be contradicted by hard empirical evidence.

§3) Disadvantages

These research methods have four disadvantages:1) The methods have limited legal precedent. 2) Attorneys and judges may be disquieted by the method's unfamiliarity. 3) To those without training in social science, the methods might seem – contrary to evidence – capriciously arbitrary. 4) Hasty readers might mistakenly think that EWs are opining on an ultimate legal issue because they use the words "commonality" "numerosity" and "typicality." However, all three terms are recognized words in the non-legal lexicon; moreover, FRE 704 allows EWs to offer evidence or opinions that "embrace[] an ultimate issue to be decided by the trier of fact" because doing so does not, in itself, usurp the court's decision-making power.

§4) Advantages

The methods have five advantages: 1) As FRE 702 requires, they necessarily utilize "sufficient data" (otherwise the reliability coefficient would expose results as insufficiently replicable); they draw on "reliable principles and methods"

---

[10] Rosenthal, R., & Rosnow, R. L. (2007). Essentials of behavioral research: Methods and data analysis (3rd ed.). Boston: McGraw-Hill.

[11] Cohen, J., Cohen, P., Aiken, L. S., & West, S. G. (2002). Applied multiple regression - correlation analysis for the behavioral sciences. Hillsdale, NJ, USA: Erlbaum Associates

with a proven track record in peer-reviewed research; and it's virtually guaranteed that they will be "applied…reliably" to the facts at bar because courts require transparent analysis methods, and EWs on both sides of the litigation must master these standardized methodologies during doctoral training in research psychology. 2) The methods satisfy all five factors considered hallmarks of scientific rigor under *Daubert*. 3) The methods have extensive precedent in other types of litigation. 4) These methods provide informative, unique, and dispositive empirical evidence that – as shown by *Robinson* – has already proven its ability to assist the court and withstand a Daubert challenge. 5) The methods have a long and distinguished history of good reliability and validity in behavioral science's peer-reviewed research.

§ 5) Summary

On balance, it's sensible and timely that attorneys and psychologists adopt these new methods. They hold considerable promise for providing indispensable unique empirical evidence that will assist the court during motions for class certification.