

Measuring the Spirit's Move: The Boon and the Bane of Empirical Methods in the Study of Evangelism, Conversion, and Spirituality.

Brian Kelly

Brian Kelly, associate professor of Practical Ministry at Southeastern University, Lakeland, Florida, earned his Ph.D. from Fuller Theological Seminary in Practical Theology, with an emphasis in spirituality/evangelism. He also is the president and founder of Barnabas Missions Unlimited, a non-profit corporation committed to informing, equipping, recruiting, and connecting lay persons, pastors, and all those committed to the task of evangelism.

Introduction

If church leaders, preachers, pollsters, and Christian educators are serious about reaching out to the more scientifically minded in our culture, they should carefully consider critical approaches to their use of certain empirical methods. This essay attempts to highlight some of the most important facets of research methods used in the study of the work of the Holy Spirit in the life of Christians.

When discussing the use of empirical methods in measuring the move of the Holy Spirit, it would be wise to begin with an understanding of what is meant by “empirical.” The idea of empirical methods is grounded in the four steps of scientific method—observation, hypothesis, experimentation, and conclusion. Empirical data gathering is critical to each of these stages in both the natural and social sciences. Empiricism can be defined as the belief that truth or knowledge is founded in what can be seen and measured with

the five senses. Scientific theories look to precise measurements of data that can be gathered and analyzed in experiments that can be replicated, tested, and even falsified if the data so indicates.

Most Westerners today readily embrace empiricism in some degree or form. The universal acceptance of empirical methods (such as the use of statistics for analyzing data) within the broader U.S. culture also permeates the church. Preachers, educators, and laypeople alike, embrace the use of empirical methods (primarily in the form of descriptive statistics), for the sake of measuring various “moves” of the Spirit within our faith communities, supporting various assertions in the pulpit and the classroom, and throughout popular, as well as academic literature. The use of such “stats” is nearly ubiquitous, thus their use necessitates critical-thinking skills in order to readily differentiate between assertions that have veracity and so-called “junk science.”

This presentation will attempt to critically analyze the widespread use of empirical methods in the measurement of such things/phenomena as conversion, evangelism, and spirituality. This discussion will address key issues of basic experimental design, operational definitions of the variable, sampling frames, and the uses and limitations of self-report measures in analyzing the work of the Spirit in the church today.

Foundational to the discussion is a more basic question of whether something as ethereal as the “move of the Spirit” can actually be measured at all. Or perhaps, should it be measured at all, or is it only a matter of intuition? Jesus seems to imply humans have an inherent dual capacity for both. “The wind blows wherever it pleases. You hear its sound, but you cannot tell where it comes from or where it is going. So it is with everyone born of the Spirit” (John 3:8). If we can measure the pneuma with our senses (hear it), then perhaps it is also part of God’s intentions for us that we can empirically measure these “moves of the Spirit” in order to build faith in our understanding of his greater and special revelation. Does the Great Horticulturalist, who knows more than anyone else about seeds and their sowing, fruit trees, vineyards, and such, not also want us to be bearers of good fruit? If he says “by their fruit you will know them” (Matthew 7:15–20), doesn’t that infer some form of measurement and assessment of that everlasting fruit he has chosen us to yield?

Since the use of empirical methods is so pervasive in the Church today, it is not really an issue of whether to use them—the flood water is over the dam, and there is no stopping it! Rather, the question is how we can best use these methods in a scientific, yet theologically judicious manner. This presentation will attempt to address this question by working through the key issues raised in this paper in a logical order based on standard experimental design.

Defining and Measuring the Spirit's Influence

The first issue of any good experimental design, from a statistics standpoint, will deal with the nature of the “operational definition” of the variables involved. How does one define conversion, evangelism, or spirituality in a given faith context?

For example, when considering conversion, there are many different theories concerning this process. In most Evangelical faith traditions, conversion is often measured by subsequent water baptism as a milestone marking genuine transformation and commitment to Christ. However, in the sacramental traditions, it may involve more of a life-long process of infant baptism, first communion, confirmation, and other such rites without any real certainty of conversion until one reaches the judgment throne of God. This is something empirically immeasurable—what theology refers to as mystery.

So, how one defines the variables in question is of critical importance in determining how best to measure them. For instance, how one defines evangelism can vary greatly based not only on the faith tradition of which one is a part, but individual experience as well. Assessing people's involvement in such things as evangelism requires assessment from a number of different perspectives, i.e., psycho-socio perspectives commonly address the cognitive, behavioral, and affective domains.

Once the researcher has settled on a credible definition of the variable in question, then a determination must be made on how best to measure it. If there are no accepted norms within the extant literature in the field of study then perhaps a more open-ended qualitative approach to the operational definitions may be in order. However, if considerable work has been done in that particular area, then credible researchers should build on the existing scholarly record and utilize measurements already developed, rather than adding more instruments to the field.

Often, a mixture of both quantitative and qualitative approaches is desirable to most accurately measure the variables in question. For instance, currently a major research project funded by the Templeton Foundation is investigating the effects of “divine love” on sacrificial service to others and its effect on revitalization within the Assemblies of God. This work uses a mixture of qualitative interviews with quantitative analysis of that data to reach conclusions regarding a variable that measures the process of interaction between the “love of divine and human actors” and its outplay within those Pentecostal churches and organizations under study.

In spite of the wide use of both descriptive and inferential statistics from broad sampling frames that attempt to survey a nationwide population, these sampling frames may be mostly irrelevant to the issues that unique

and anomalous individuals are facing on a day-to-day basis. In these cases, perhaps individual longitudinal studies are much more applicable than broad generalizations extrapolated from scientific random samples. Issues of individual conversion and/or spiritual development would likely best be investigated through this kind of methodology.

Sampling Frames

Once one has accurately and operationally defined the variables, a sampling frame then needs to be selected. How this is done can greatly influence the outcome. Pollsters rely on the idea of null hypothesis significance theory in statistics to arrive at 95 percent confidence intervals. The idea is that if you conduct a “scientific” or “simple random” sample, there is a 95 percent likelihood that every person in that population (for example, the U.S.A.) regardless of race, age, gender, geographical location, faith tradition, etc., could have been asked the question. If that is true, then the statistical result at an N of 2,050 would be no different than if the N were 205,000 or 2.5 million.

The reality is that few researchers beyond professional pollsters like Barna and Gallup have the resources to do such scientific random sampling. Rather, most research of religiosity and/or spirituality is done with “haphazard” or “convenience” samples. The respondents are contacted through catch-as-catch-can approaches and these are inherently limited. Yet, readers of such research will automatically generalize these findings to broader populations (the entire U.S., if you will) even when the sample sizes are really quite small and non-representative. For instance, researcher Thom Rainer polled 306 people who are not church attendees or believers and extrapolated those findings to the entire U.S. population.

Choices regarding who to poll can have dramatic impact on responses because of the nature of the respondents. For instance, when polling unchurched persons about their attitudes regarding the church in general, some stark differences in opinions emerge based on just the sampling frame. Kimball, in his book *They Like Jesus But Not the Church*, reported on attitudes toward the church from those outside the church. His findings came from acquaintances he polled in Santa Cruz, CA, who were young adults and were not likely ever to be a part of a church.

Rainer, for research in *Surprising Insight From the Formerly Unchurched* had a much different approach. He polled 353 people who “for all or a large portion of their lives were not in church” but now are Christians attending church. He surveyed those who are now a part of a local church on why they became Christians. The construct validity of this experimental design is much more reliable. It addresses some of the same issues as Kimball’s, but draws data from a much more telling sample.

Kinnamman, President of the Barna Research Group, in his popular book *UnChristian*, uses scientific, random sampling techniques familiar to those employed in Null Hypothesis Significance Testing (NHST) to arrive at his conclusions. While the sampling frames are valid, the operational definition of what constitutes a Christian is problematic. His basic premise that perception is reality in the marketplace of ideas in U.S. culture can be affirmed to some extent. He found that 84 percent of his respondents, who were young people between the ages of 13–29 years, know a Christian, but only 15 percent see a difference. Perhaps this is indicative of a significant design assumption—that how one defines a Christian is largely left to self-identification.

In regard to such self-descriptions, teenagers at the end of the last century overwhelmingly self-identified as Christians. “For more than a decade, regardless of their beliefs and church attendance, more than four out of five teens have been describing their faith affiliation as Christian. The last few years, however, have brought about a minor decline in such self-descriptions. While 88% called themselves Christian in 1997, the figure dropped to 84% in 1998 and just 82% in 1999.” Can we really accept as fact that 8 out of 10 young people in our country are Christian?

Barna uses two distinct definitions for Christians. One for born-again Christians and another for Evangelical Christians. Among the so-called self-identified born-again Christians polled for the *UnChristian* study, 59 percent of born-again Busters (ages 23–41) think cohabitation is morally acceptable compared to 33 percent of born-again older adults (age 42+). Consistently, 44 percent of those same Busters and 23 percent of the older cohort see sex outside of marriage also as morally acceptable. Nearly a third (28 percent) of those Busters thought having a sexual relationship with someone of the same sex is morally acceptable, and even 13 percent of the older adults agreed. If this data is indicative of all who self-identify as born-again Christians, no wonder the culture at large does not see any difference!

How Valid Are Our Statistics?

Not only are there problems with sampling biases and operational definitions of variables under study, but statistics are commonly used across Christianity to support validity of ministries, outreaches, and revivals. The accepted stamp of authenticity for true revival is the number of genuine conversions.

The so-called “Pensacola Portal” or revival at Brownsville, FL AG provides a case in point. This researcher was familiar with three very disparate sources and data concerning the number of conversions associated with the “Spirit’s outpouring” there. Dr. Michael Brown, who was in charge of the

discipleship ministries at Brownsville for a time, lectured at Fuller Seminary and recounted from their own periodical that they had recorded over 115,000 conversions since the revival had started. In the same period the Los Angeles Times reported between 175–190,000 converts at Brownsville. Contemporaneously, one TBN evangelist, Rev. Rodney Howard Brown, in an interview with a television host reported he had heard that over 300,000 people had come to Christ there since the revival's origin.

When asked about the growth of the church, Dr. Brown reported that church attendance had grown from slightly over 350 to more than 1,500 in regular weekly attendance. This information brought a question by someone in the audience: "What happened to the other 113,500 believers whose conversion they had recorded?"

The issue is how they gathered the data, and how they defined conversion. Every evangelist knows the challenges in collecting such data, for example. People often come forward to "altar calls" for prayer for a variety of things. Most evangelists I've heard are not precise enough about what people are being asked to respond to by coming forward. The result is murky data regarding the number of genuine first-time conversions, hence the widespread use of a common euphemism among clergy—"evangelistically speaking," wink wink. The implication is that the data itself is as elastic as the claimant wants to make it. Interestingly, it may be the pastors themselves who are "fudging." One of the more striking comparisons is the large number of conversions recorded over the last decade in the Assemblies of God according to the pastor-reported Annual Church Ministries Report, with little or no actual growth in church attendance reported as a result of all these conversions.

Another common myth I hear often in the body of Christ is that you can make statistics say anything you want. This is true only if you don't have a commitment to report things in an ethical manner and in a way that is credible to critical thinkers.

Advantages and Limitations of Self-Report Measures

Faking Good

Indicative of a prevalent response set known as "faking good," Americans overwhelmingly identify themselves as Christians but apparently do not practice what they preach. Rainer reports that over 80 percent of those who were not antagonistic or resistant to Christianity would be interested in a Bible study or learning more about Jesus. Over 90 percent of those same respondents say they would attend church if someone invited them. However, responding this way to a researcher who asks such things over the phone is very different from being asked directly, "Would you like to attend

a Bible study in my home or attend my church with me this week?" In the American religious psyche, there is a decided disconnect between espoused belief and actual practice.

Furthermore, church attendance figures are even sketchier due to self-report biases and how the data is gathered. Some Christian researchers, after looking at data gathered and reported by Barna and Gallup that 40 percent of Americans attend church on Sunday, contend that the number is likely only half that because of the way the data is gathered—through self-report measures. For instance, when researchers polled a congregation in Ohio about their Sunday School attendance for the prior Sunday, 60 percent said they had attended when only 38 percent had actually done so. Pastors and church leaders, who are often promised participation by their congregants at various events or programs only to find that a much smaller number of volunteers actually show up, are not likely to be surprised by such findings.

Survey Questions Influence Responses

Since one of the primary ways data is collected in the church growth/health literature is through self-report surveys, we must pay careful attention to how these self-report instruments are structured. One of the more interesting studies found that question content, format, and ordering conditions have a profound impact on responses. As an example, when respondents were asked to rate "how successful they have been in life" on a scale with semantic anchors of "not at all successful" to "extremely successful," results varied considerably depending on how the question was framed. On a scale from -5 to +5, fully 34 percent rated themselves between -5 and 0, but only 13 percent endorsed the formally equivalent score on a scale from 0 to 10. Schwarz found that although researchers tend to see questionnaires as "measurement instruments" that elicit information from respondents, a significant truth is that they are also a "source of information that respondents tend to draw on in order to determine their task and to arrive at a useful and informative answer." In other words, due to the respondents' desire to do well, the way the questions are asked, the potential answers that are offered, and even the order of the questions can have a profound impact on the results.

Advantages of Self-Report Measures

The challenge for researchers in regard to self-report measures is that in some cases they are the most efficient way to gather data from a large number of respondents. While focus group interviews are one way of assessing spirituality, individual responses to surveys on spirituality remain perhaps the most viable way of assessing larger groups over a longer period

of time. For example, Southeastern University uses a Spiritual Assessment Inventory to evaluate their College of Religion students on two domains—Awareness of God, and Quality of Relationship with God. A pre-test measure is administered on declaration of the major, which is compared later to the same post-test assessment on graduation.

Perhaps one of the more popular self-report measures is the questionnaire compiled by Schwarz. His organization, Natural Church Development, gathered data from over 65,000 churches in more than 83 countries of the world, 35,000 of these churches in the U.S. alone. The senior pastor and 30 key lay people selected from the congregation are polled. Data analysis provides churches with not only an internal comparison of the eight quality characteristics of church health, it also offers comparisons between growing, healthy, unhealthy, and not growing churches on these eight factors.

Data from these surveys provides the local church with a valid “snapshot” of where they are by comparing and helping them identify the minimum factor from among these characteristics in order to focus improvement on that factor. The results also provide empirical data for the local church leaders so basic growth/health principles can then be individualized for the local church.

Conclusion

The use of empirical methods in the study of the Holy Spirit’s work in the local church is here to stay. The challenge is to critically engage the use of such methods in a positive way. Through the careful analysis of the experimental design used in research, addressing such things as the precise operational definition of the variables, the nature of those who were sampled, how they were sampled, and a wariness to accept every assertion that is made through the use of statistical data, perhaps we can have confidence that the fruit we are claiming is indeed fruit that will last.

Notes

1. Paul C. Cozby, *Methods in Behavioral Research*, 9th ed. (New York: McGraw Hill, 2006), 4. It is also defined as “the practice of relying on observation and experiment, especially in the natural sciences.” See <http://www.merriam-webster.com/dictionary/empiricism> (accessed June 12, 2011).

2. All Scripture is from the New International Version unless otherwise noted.

3. Cozby, 68. An operational definition is “a definition of the variable in terms of the operations or techniques the researcher uses to measure or manipulate it.”

4. For a theological approach, see Richard Peace, *Conversion Paradigms in the New Testament: The Conversion of Paul and the Twelve* (Grand Rapids: Eerdmans Publishing Co., 1999); sociological, see Lewis R. Rambo, *Understanding Religious*

Conversion (New Haven: Yale University Press, 1993); psychological, see Cedric B. Johnson and H. Newton Maloney, *Christian Conversion: Biblical and Psychological Perspectives* (Grand Rapids: Zondervan, 1982); and missiological, Paul G. Heibert, "Conversion, Culture and Cognitive Categories," *Gospel in Context 1*, no. 4 (October 1978): 24-29.

5. 1 Corinthians 3:12–15.

6. Richard Gorsuch, "Measurement: The Boon and The Bane of Investigating Religion," *American Psychologist*, 39, no.3 (1984), 234. Gorsuch argues against the development of new measures when existing ones are more than adequate.

7. Michael Q. Patton, *Qualitative Research & Evaluation Methods*, 3rd. ed. (Thousand Oaks, CA: Sage Publications, 2002), 585. Patton contends that the old debate between advocates of either quantitative or qualitative methods "has softened. A consensus has gradually emerged that the important challenge is to appropriately match methods to purposes, questions, and issues, and not universally advocate any single methodological approach for all inquiry situations."

8. "Margaret Poloma on Pentecostalism, the Assemblies of God, and Godly Love" from *Research on Religion* podcast, Dec. 6, 2010, cf. <http://www.researchonreligion.org/protestantism/margaret-poloma-on-pentecostalism-the-assemblies-of-god-and-godly-love> (accessed June 12, 2011).

9. Interview with Dr. Peter Althouse on June 12, 2011, Southeastern University, Lakeland, FL. Althouse is one of the researchers involved in the study with Poloma.

10. David L. and Robin K Morgan, "Single Participant Research Design: Bringing Science to Managed Care," *American Psychologist* (2001): 119–127. The researchers note that "at an applied level, it is seldom the case that treatment objectives center on how group averages responded to manipulated variables and interventions. For parents, teachers, therapists, and others charged with changing behavior, the individual ordinarily constitutes the unit of analysis, and change makes itself known only through multiple measures taken over prolonged observational periods," 122.

11. Thom S. Rainer, *The Unchurched Next Door: Understanding Faith Stages As Keys to Sharing Your Faith* (Grand Rapids: Zondervan, 2003), 18. In Rainer's defense he argues that his sample "represented every conceivable demographic area of America."

12. Dan Kimball, *They Like Jesus but Not the Church: Insights from Emerging Generations* (Grand Rapids: Zondervan, 2007).

13. Thom S. Rainer, *Surprising Insights from the Unchurched and Proven Ways to Reach Them* (Grand Rapids: Zondervan, 2001).

14. *Ibid.*, 22.

15. David Kinnaman and Gabe Lyons, *Un-Christian: What a New Generation Really Thinks About Christianity and Why It Matters* (Grand Rapids: Baker Books, 2007). See page 250 for his methodology, which was an aggregate of numerous studies performed by the Barna Group from 1995–2007, with sample sizes ranging from a qualitative study in a smaller focus group (n = 27), to much larger sampling frames, both scientific and random (ranges from n = 613 to n = 1020) and self-select online polls (from n = 102 to n = 2409).

16. *Ibid.*, 48.

17. George Barna, *Third Millennium Teens: Research on the Minds, Hearts and Souls of America's Teenagers* (Ventura, CA: Barna Research Group, 1999).

18. *Ibid.*, 247. Born-again Christians are defined as people who say they have made a personal commitment to Jesus Christ that is still important in their life today, believe they will go to heaven when they die because of confession of sin and acceptance of Christ as savior, although they are not asked directly if they are born again. Evangelicals meet the born-again criteria plus 7 other conditions: (1) saying their faith is very important in their life today; (2) believing they have a personal responsibility to share their religious beliefs about Christ with non-Christians; (3) believing that Satan exists; (4) believing that eternal salvation is possible only through grace, not works; (5) believing that Jesus Christ lived a sinless life on earth; (6) asserting that the Bible is accurate in all that it teaches; (7) describing God as the all-knowing, all-powerful, perfect deity who created the universe and still rules it today.

19. Lecture Session with Dr. Michael Brown, Nov. 1997, Fuller Theological Seminary, Pasadena, CA. See "Pentecostal Church Shakes Things Up Along Gulf Coast," *Los Angeles Times*, May 3, 1997, 19; and personal recollection of a TBN broadcast.

20. Recent data showed an increase of 11.3% (n = 182,675) from 1998 to 2008 in AG U.S. Sunday Morning/Major Worship Service Attendance, yet the number of conversions reported every year for that 10-year period were over 400,000 annually. See Office of the Statistician, General Council of the Assemblies of God, Springfield, MO.

21. According to Gallup's website "in 2003, 42% of U.S. adults said they were born again or evangelical; the 2004 percentage is 41%." See Albert N. Winseman, "Who Has Been Born-Again?" <http://www.gallup.com/poll/14632/Who-Has-Been-Born-Again.aspx> (accessed June 12, 2011).

22. Rainer, *Unchurched Next Door?*

23. C. Kirk Hadaway and Penny Long Marler, "How Many Americans Attend Worship Services Each Week: An Alternative Approach to Measurement," *Journal for the Scientific Study of Religion* 44, no. 3 (2005): 307-322; and C. Kirk Hadaway, Penny Long Marler, and Mark Chaves, "Over-reporting Church Attendance: Evidence that Demands The Same Verdict," *American Sociological Review* 63, no. 1 (1998): 122-130.

24. Hadaway, Marler and Chaves, 127.

25. Norbert Schwarz, "Self-Reports: How the Questions Shape the Answers," *American Psychologist* (1999): 93-105.

26. *Ibid.*, 95.

27. *Ibid.*, 103.

28. Mark Cartledge, *Testimony in the Spirit: Rescripting Ordinary Pentecostal Theology* (Burlington, VT: Ashgate Publishing, 2010). Cartledge uses this technique of recording group interviews then mining the transcriptions for insightful data.

29. For an exploratory and confirmatory factor analysis of these domains with other common psychometrics, see Todd W. Hall and Keith J. Edwards, "The Spiritual Assessment Inventory: A Theistic Model and Measure for Assessing Spiritual Development," *Journal for Scientific Study of Religion* 41, no. 2 (2002): 341-357.

30. Christian A. Schwarz, *Color Your World with Natural Church Development: Experiencing All That God Has Designed You to Be* (St. Charles, IL: ChurchSmart Resources, 2005).

31. *Ibid.* The eight characteristics are defined as Empowering Leadership, Gift-Oriented Ministry, Passionate Spirituality, Functional Structures, Inspiring Worship Service, Holistic Small Groups, Need-Oriented Evangelism, and Loving Relationships. The emphasis is put on the adjective as the primary operational definition.