

*Ghost in the Analysis: The Good, the Bad, and the Ugly in the Use of Empirical Methods
In the Study of the Holy Spirit's Influences in the Church Today*

Brian M. Kelly, Ph.D.

Southeastern University

Practical Theology Interest Group

**Presented at the 41st Annual Meeting of the Society of Pentecostal Studies, Practical
Theology Interest Group at Regent University, Virginia Beach, VA Spring 2012.**

Introduction

How many times have grandiose claims been made of huge numbers of conversions as a stamp of authenticity regarding a “revival” somewhere in the world?¹ Testimonies of Christian conversion are rightly seen as one of the badges of authenticity that earmark a genuine move of the Spirit among humans.² Yet, too often unsubstantiated claims of conversion, or other kinds of manifestations of the Spirit, are made without much critique regarding these claims. Can we accurately measure the Spirit’s influence upon humanity? Or some might ask, should we? Can we be more precise in our measurement of what God is doing in the world today without embracing a wholesale empiricism that denies divine intervention is even a possibility in the world it seeks to analyze?

This research will attempt to critically analyze the widespread use of empirical methods in the measurement of the Holy Spirit’s interaction with humans in conversion, evangelism, and spirituality. The widespread use of statistics by researchers, church leaders, educators, and lay people frequently is without critical discernment. Basic errors in fundamental research design, questionable data analysis and/or presentation of results, and dubious conclusions are often accepted uncritically because of their implied sophistication. Unfortunately, fundamental concepts concerning research methods such as; the operational definitions of variables, careful choice of sampling frames, interviewer and observer biases, and the uses and limitations of self-report measures seem to be foreign to many when analyzing the work of the Spirit in the church today.

¹ The so-called “Pensacola Portal” or revival at Brownsville Assembly of God in Pensacola, Florida provides a case in point, but such practice is replete in Christendom today. This is addressed in more detail in the operational definitions section of this document.

² For example the Azusa Street Mission was known to be “aggressively evangelistic” and grew from a small prayer meeting of 15 people into an “internationally acclaimed congregation of hundreds in just three months,” Cecil M. Roebeck, Jr. *The Azusa Street Mission and Revival* (Nashville; Thomas Nelson, 2006), 6.

This research will attempt to lay out the appropriate means and measures that should be employed, the abuses that are pervasive, and call attention to pitfalls that should be avoided at all costs. It will address the widespread use of both quantitative and qualitative methods in empirical research especially as they pertain to those topics. This will be done by illustrating what is considered acceptable by the best practitioners (the good), the challenges that lead to abuse of these methods (the bad), and some of the more outlandish examples of outright neglect of critical methods when available (the ugly).

Measuring the Spirit's Move: Can We? Should We?

Use of Empirical Methods

The discussion concerning the use of empirical methods in measuring the move of the Holy Spirit begins with an understanding of what is meant by “empirical.” Empirical methods are grounded in the 4 steps of scientific theory – observation, hypothesis, experimentation, and conclusion.³ Empirical data gathering is critical to each of these stages in both the “hard” natural and “soft” social sciences. Empiricism can be defined as the belief that truth or knowledge is founded in what can be seen and measured with the 5 senses.⁴ Scientific theories look to precise measurements of data that can be gathered and analyzed in experiments that can be replicated, tested, and even falsified.

Empirical methods are often integrated into theological research – especially in the field of practical or applied theology. Perhaps nowhere is this use more prevalent than in the study of

³ This approach is mirrored in practical theology with 4 phases of dynamic interaction; 1) Specific Description, 2) Developing Dialogue, 3) Reasonable Intentions, and 4) Sensible Applications, abridged from Don S. Browning, *A Fundamental Practical Theology* (Minneapolis: Fortress Press, 1991), 58.

⁴ 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 January 15, 2012).

the psychology of religion where researchers interested in Christian conversion have spent considerable effort to gather data using empirical methods.⁵

Most Westerners today readily embrace empiricism in some degree or form, e.g., the universal acceptance of empirical methods (such as the use of statistics for analyzing data) that permeates the Christian church. Preachers, educators, and lay people 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. Whether in the pulpit, classroom, throughout popular or academic literature, the use of such “stats” is nearly global. Such use makes it imperative that Christians develop critical thinking skills to readily differentiate between assertions which have veracity and what some have coined “junk science.”

Perhaps a more fundamental question might be whether something as ethereal as the “move of the Spirit” can *actually* be measured due to the limitations of natural science. In addition to the question of whether the Spirit’s influence can be measured is a more philosophical question of *should* it be measured at all, or is it only a matter of intuition? One religious research argues the use of empirical research in the study of religion “allows one to test his decisions in the crucible of experience and obtain a more accurate analysis than can be gained from casual observation.”⁶ Simply “flying be the seat of your Pentecostal pants”⁷ when it comes to important assessments of what God is doing in the midst of a situation is even more detrimental.

Many Pentecostals might decry the use of empirical studies of human behavior to

⁵ See the Operational Definitions section of this paper for a more comprehensive literature review regarding the topic of conversion.

⁶ Richard L. Gorsuch, “Research and Evaluation: Their Role in Decision Making in the Religious Setting,” *Review of Religious Research*, Vol.17, No. 2 (Winter 1976), 94.

⁷ A phrase I like to use in my classrooms to depict haphazard, intuitive, emotive and most often erroneous decision making based on a so-called “leading of the Spirit.”

measure the Spirit's move as lacking the kind of intuitive spirituality so highly prized by those led by the Spirit. But I believe there is a place for the "Ghost" in the analysis as well. For, "if the Holy Spirit can lead us when we are unprepared how much more can he lead us when we are prepared."⁸ Pentecostals would do well to consider the fact that setting and measuring certain behavioral objectives and criteria might seem limiting on its face, but failure to do so is to trust inherently in "untutored" and so-called "expert" opinion which has been proven to be more often erroneous than not.⁹ "While the behavioral objective approach can be criticized in practice, almost every criticism of an instrument-oriented approach applies with even greater force to approaches without behavioral objectives."¹⁰

Jesus' words imply humans have an inherent dual capacity for acquiring both intuitive and empirical knowledge. "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."¹¹ If we can measure the *pneuma* with our senses (hear it) 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 revelation. Does not the Great Horticulturalist, who knows more than anyone about seeds and their sowing, fruit trees, vineyards and such, also want us to be bearers of good fruit? If he says "by their fruit you will know them" doesn't that infer some form of measurement and assessment of that everlasting fruit he has chosen us to yield?¹²

The use of empirical methods is so pervasive in the Church today it is not really an issue whether to use them – the flood is over the dam and there is no stopping it! Rather, the question is how we can best use these methods in scientific yet theologically judicious ways, for empirical

⁸ I attribute this quote to Rev. David Jones my Bible professor and pastor at Trinity Bible College in Ellendale, ND.

⁹ Gorsuch, 97.

¹⁰ *Ibid.*, 95.

¹¹ John 3:8. All Scripture is from the *New International Version*, unless otherwise noted.

¹² Mt. 7:15-20.

methods carry with them inherent bias. In practical theology there is a need to look critically at the experiential dimensions of faith, what has been called “ecclesiastical statistics” which in contrast to philosophical or theological theories is an “empirical science related to perceptible reality.”¹³

The Reflective Practitioner

However, as already noted, empirical studies are not without their limitations which will be discussed in detail in the following pages. Unlike some fields of study (hermeneutics, homiletics, systematic theology) there is no distinctly “Christian” or “Pentecostal” way of doing empirical studies. Rather, one must trust in the presence of the Spirit (“ghost in the analysis” again) within the researcher to help guide the research.¹⁴ Those practitioners who are working in the field are in the best position to help gather and integrate information in the process of research and evaluation.¹⁵

But such integration comes with a price and the Spirit must help us to discern what is true and what is not.¹⁶ Furthermore, a full recognition of the spiritual dimension in the human condition is not merely the product of intellect. Beyond rational explanation, it exceeds mere cognition to touch the affective and intuitional dimensions.¹⁷

We have two poles really. On one side is the widespread use of phenomenological methods which are often suspected of “ignoring or even contradicting common rules and

¹³ Hans-Günter Heimbrock, “From Data to Theory: Elements of Methodology in Empirical Phenomenological Research in Practical Theology,” *International Journal of Practical Theology*, vol. 4, (2005), 274 .

¹⁴ Gorsuch argues that religious practitioners are the best researchers in the field of research for the church because “the gathering of and utilization of information by empirical means must be guided by and integrated into a deep understanding of the area” to be studied, see Gorsuch, 100.

¹⁵ “Since they are the ones who will ultimately use it, the information must meet their standards and be integrated into their theoretical frameworks,” Ibid.

¹⁶ John 16:13: “But when he, the Spirit of truth, comes, he will guide you into all truth.”

¹⁷ An extensive study of over 4,000 people concerning the nature of human spirituality concluded that the main characteristics of religious and spiritual experiences were illustrated in subject's "feelings for a transcendental reality" and provided empirical support for the significance the human conscience places on the reality of the spiritual dimension beyond the individual self, or as the author stated a "sense of presence (not human)," Alastair Hardy, *The Spiritual Nature of Man* (Oxford: Clarendon Press, 1979), 131.

standards of empirical research, such as objectivity and generalization,” because such methods are not committed to the formulation and subsequent testing of “exact hypotheses about distinct objects.”¹⁸

On the other side, empirically-based theories do not, as some might assume, deal strictly with facts. For there is no such thing as a *tabula rasa*, or blank slate of the mind, rather all research “inevitably bases its propositions upon large theoretical concepts, such as reality, life, and religion. Even the so-called starting points of data, empirical findings, or phenomena carry a heavy theoretical load.”¹⁹

The answer is a balance between the two, based on an ethos led by the Spirit of Christ. After all, if Jesus Christ is King of Kings, Lord of Lords and Master of the Universe then it makes sense that he is also the most adept and consummate of researchers as well, “the most intelligent person who ever lived on earth.”²⁰ Perhaps the ideal is the Spirit-led “reflective practitioner” – a person who strives to combine the best of all these attributes; empirical research - with the practical wisdom to implement it. By prayerful consideration and exercise of a Holy-Spirit-baptized intellect, Pentecostal researchers can trust that their research designs, data gathering, analysis, and presentation will be the best that God can produce through a frail humanity immersed in a fallen world.

¹⁸ Heimbrock, 277.

¹⁹ Ibid., 280.

²⁰ Dallas Willard has asserted that “we need to understand that Jesus is a *thinker*, that is not a dirty word but an essential work, and his other attributes do not preclude thought, but only ensure that he is certainly the greatest thinker of the human race...” Dallas Willard, *The Great Omission* (New York: HarperCollins, 2006), 189.

**Defining and Measuring the Spirit's Influence - Key Concepts for Critical Analysis –
Sound Research or Junk Science?**

Operational Definitions of the Variable

The Good

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.²¹ How does one define conversion in a given faith context? Or evangelism? Or spirituality?

For example, when it comes to conversion there are many different theories concerning this process.²² 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.²³ Assurance of final conversion is thus relegated to the hereafter, 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, as alluded to, how one defines conversion 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

²¹ An operational definition is “a definition of the variable in terms of the operations or techniques the researcher uses to measure or manipulate it,” cf., Cozby, 68.

²² 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*, Oct. 1978.

²³ 1 Cor. 3:12-15.

of different perspectives, (i.e., psycho-socio perspectives commonly address the cognitive, behavioral, and affective domains). In fact, the idea of defining what it means to be a Christian has been the focus of volumes of Christian literature that address spiritual formation, discipleship, ethics, and a whole host of other doctrinal issues related to Christian spirituality.

This issue of definitions is an important one especially for Pentecostals, as there is considerable debate within the movement itself about what it is that constitutes a Pentecostal.²⁴ The present consensus seems to be that it is more a matter of self-identification with a Pentecostal church, denomination, or set of beliefs and practices than it is a matter of an experience of Spirit baptism.²⁵ This paper will not attempt to enter the theological fray in that regard. The topic is raised to illustrate how important the matter of precise nomenclature is even within theological circles.

Defining Conversion as an Example

As another poignant example of how important operational definitions can be, one need only look further at the topic of conversion in theological, sociological, and missiological literature to find a wide disparity across the board. Furthermore, how this gets defined in the local church is a matter of much doctrinal controversy.²⁶ When this researcher completed his dissertation an extensive review of the contemporary literature concerning numerous empirical

²⁴ Just attend any Society for Pentecostal Studies meeting and raise the issue and see what kind of robust discussion results.

²⁵ As Anderson notes “the term Pentecostal is appropriate for describing globally all churches and movements that emphasize the working of the gifts of the Spirit, both on phenomenological and on theological grounds – although not without qualifications,” see Allan Anderson, *An Introduction to Pentecostalism* (Cambridge: Cambridge University Press, 2004), 13. For a more thorough analysis of the discussion see *Ibid.*, 9-15, Amos Yong, *The Spirit Poured Out on All Flesh* (Grand Rapids: Baker Academic, 2005), 18-22, and Frank D. Macchia, *Baptized in the Spirit* (Grand Rapids: Zondervan, 2006), 47-48.

²⁶ Witness the stark differences between the Sacramental view of conversion (especially the efficacy of infant baptism) and Evangelicalism’s so-called “believer’s baptism.”

studies of conversion and evangelism efforts was performed.²⁷ If the search is limited to those studies that honored a scientific approach, rather than a purely theoretical or theological framework, it becomes apparent that the most comprehensive and valid empirical studies are to be found primarily within the psychological field.²⁸ Stretching back more than a hundred years there is an extensive history of investigation concerning the various factors surrounding conversion within the American context.

As early as 1896 Starbuck completed an extensive, three year qualitative study of conversion which primarily focused on those who experienced a "sudden awakening."²⁹ His approach was to set the standard in the study of conversion for nearly a half century. Not only did Starbuck attempt to appraise the factors preceding, during, and subsequent to the conversion experience, but his was probably the first study to chart the age of conversion and thus illustrate the correlations between puberty and conversion.³⁰

Thirty years later, Elmer Clark was to complete another comprehensive study which involved some 22 open-ended questions concerning conversion.³¹ Clark argued for the validity of the use of questionnaires as a means of surveying the "personal testimony of experience" on

²⁷ Brian Kelly, *Toward A Practical Theology of Evangelism: Spiritual Journey as an essential Paradigm for the Conversion Process* (Ph.D., diss., Fuller Theological Seminary, 1999).

²⁸ Some in the church would object to attempts to limit the study of faith to empirical methods alone, leaving out the metaphysical and subjective elements. As W.H. Clark noted about conversion "the psychologist is not competent to define the place of God in the transaction. The spiritual significance comes not from the psychological experience but the interpretation of it," cf. Walter Houston Clark. *The Psychology of Religion* (New York: Macmillan Press, 1958), 200. He goes on to point out that God is inscrutable, "hence it is hardly likely that a finite psychologist with merely human powers of comprehension will be able to pronounce on the ways of an infinite God beyond what he is able to observe. The psychologist, as a *psychologist*, can neither include nor exclude the influence of God in the conversion process." *Ibid.*, 200-201.

²⁹ Due to the nature of Starbuck's methodology, using lengthy open-ended questions in his survey, he found it difficult to separate the two groups into what he termed "sudden awakenings and gradual growth." Only those with a "marked change" were put in the former class, and of these, he focused only on those responses that were "sufficiently complete," numbering some 137 subjects. Cf., Edwin Diller Starbuck. "A Study of Conversion," in *The American Journal of Psychology*, Vol.8, (1897): 270.

³⁰ He noted that "the years of greatest frequency of conversions correspond with periods of greatest bodily growth for both sexes," and that "there is a correspondence between the periods of most frequent conversions and puberty in both sexes," *Ibid.*, 271- 272.

³¹ Elmer T. Clark. *The Psychology of Religious Awakening* (New York: Macmillan Press, 1928), 24-25.

the basis of its acceptance between the scholars and the external tests of validity that were applied in his day.³² Like others before him, Clark acknowledged the difficulty of establishing fixed categories of spiritual experience.³³ He grappled with the whole definition of religious conversion narrowing it to those "emotional experiences of a more radical sort in which a sudden change from irreligion or non-religion to religion is involved."³⁴

In spite of the definitional ambiguities inherent to such studies, Clark's work was helpful in that it expanded the categories concerning his subject's experiences into three distinct categories: 1) the "definite crisis awakening" 2) the "emotional stimulus awakening" and 3) the "gradual awakening."³⁵ His expansion of conversion into three categories of experience proved helpful in delineating the difference between various kinds of experiences and provided foundational statistics for the study of the process.³⁶

Another twenty years would pass until following World War II, Gordon Allport and his colleagues set out to test the religion of postwar college students by polling 500 students enrolled at Harvard and Radcliffe College some fifteen months after the war had ended.³⁷

Ten years later in 1958, one of Allport's protégés, Walter. H. Clark would reiterate the concern of prior researchers who dealt with self-report measures of conversion. He noted that in

³² Ibid., 30.

³³ "So prominent is the personal equation that it may be said that there are as many types of religious awakenings as there are religious individuals," Ibid., 34.

³⁴ Ibid., 36, *contra* Starbuck who saw it as primarily an "unselfing," where the individual "learns to transfer himself [*sic*] from a center of self-activity into an organ of revelation of universal being." This involved a process of conviction, crisis and new life, cf., Starbuck, 303-305.

³⁵ They are defined respectfully as 1) "a definite crisis is reached and passed and an actual change of attitude is effected;" 2) "no special change is effected, but an event serves as a stimulus to awaken religious consciousness;" 3) "the religious life flows onward like a stream, enlarging and growing but striking no obstructions and forming no cataracts," cf., E.T. Clark, 39f.

³⁶ Clark's data showed that among the three types of awakenings among the total of 2174 persons polled, 66.1% had a Gradual conversion, 27.2% claimed it was the result of Emotional Stimulus, and 6.7% saw it as the result of a Definite Crisis, cf., Ibid. 48, 86.

³⁷ Allport's findings concerning the age or "types of subjective religious awakening" did not vary significantly from E.T. Clark's earlier findings. See Gordon W. Allport, *The Individual and His Religion: A Psychological Interpretation* (New York: Macmillan Company, 1950).

most analyses "little distinction is made between the conversion that has occurred on the spur of the moment in the gospel tent and that which is the culmination of years of inward struggle."³⁸

He also made a salient observation concerning E.T. Clark's findings concerning the prevalence of the crisis conversion among religious workers,³⁹ noting that this observation "should give some pause to those advocates of gradual growth who at the same time are concerned about the recruitment of religious workers."⁴⁰

At that point in time, the development of two primary categories for determining the duration of time for conversion to occur was prevalent: sudden and gradual. This is reminiscent of James' description of conversion as the unification of the self, categories he described as once or twice-born.⁴¹ In the historical development the sudden category was further divided into those kinds of experiences evoked either from crisis or due to emotional stimuli over a longer period, but still resulting in a sudden and marked change. The gradual category up to that time was without nuances.⁴²

In the late 60's Geoffrey Scobie added a further dimension (what he considered a third), to the conversion process that he described as the "unconscious conversion."⁴³ This category was for those who were brought up in a Christian home and included those respondents who could not think of a time in their lives when they were not a Christian, when they did not accept the

³⁸ W.H. Clark, 213.

³⁹ 17.7% of these workers reported the crisis conversion compared with 2.8% for non-religious workers or to a degree more than six times more prevalent, see E. T. Clark, 112-115.

⁴⁰ W. H. Clark, 214.

⁴¹ William James, *The Varieties of Religious Experience*, Longmans, Green, and Co., 1902; repr., (New York: Penguin Books, 1985), 166-88. James saw the process of unification as either gradual or abrupt, *lysis* or *crisis*, see *Ibid*, 175-83.

⁴² The characteristics of the gradual conversion were seen as: 1) an active search for meaning and purpose; 2) the absence of emotional crisis or feelings of guilt or sin (almost purely cognitive); and 3) a continual and progressive deepening of the faith which respondents have cognitively assented to accept, cf. Bernard Spilka, Ralph W. Hood and Richard L. Gorsuch. *The Psychology of Religion: An Empirical Approach* (Englewood Cliffs, N.J.: Prentice Hall, Inc. 1985), 204.

⁴³ Geoffrey E. W. Scobie. *Psychology of Religion* (New York: John Wiley and Sons, 1975), 50.

Christian belief.⁴⁴ In other words, they would likely respond that they “have always known the Lord.”

Scobie's assertions are indicative of the historical shift that occurred in the study of conversion. The focus was shifting from a concept of conversion as "a totally individual phenomenon, simply defined, to one showing many possible forms and expressions."⁴⁵ Where the original focus was on "what stimulated and followed conversion, in contrast, contemporary workers are emphasizing the *process* of conversion."⁴⁶

As time progressed, the studies became more and more focused on the causes and results of conversion rather than the types of conversions. Paloutzian and his co-authors reaffirmed that it was useful to classify Christians according to belief and type of conversion experience.⁴⁷ Similar to Scobie, they also used three different types of conversion in their study.⁴⁸

Finally, Christine Liu, believes the decision point to be more influential than the process.⁴⁹ She found that "individuals that made a conscious decision to be a Christian were more religious than those who had not made a conscious decision" and that no significant differences in religious devotion were found between those whose conscious decisions were

⁴⁴ He found that 30% of his sample of 170 theological students claimed to have followed this pathway to religious belief. In his study only 20 percent claimed a sudden conversion, with the 30% above, this left 50% who had what he termed a “gradual” conversion, that is where the process of growth of belief extends over a period of time, days, months or even years. During this period the person moves from a position where they are rejecting Christianity as a whole, or some specific part of it, to a point where this rejection has changed to acceptance, *Ibid.*, 50-51.

⁴⁵ Spilka, et al., 223.

⁴⁶ *Ibid.*

⁴⁷ Raymond F. Paloutzian, Steven L. Jackson and James E. Crandall. "Conversion Experience, Belief System, and Personal and Ethical Attitudes" in *Journal of Psychology and Theology*, vol. 6 (Fall, 1978), 266. “Different types of conversion may effect intrinsicness, meaningfulness, and maturity of the religious commitment,” *Ibid.*, 267.

⁴⁸ They performed two studies one among college students and another among older adults. A composite of the two studies showed that 49% had a “Unconscious” conversion, 33% reported Gradual, and 18% a Sudden experience (this was considerably higher in the adults than the students 23% versus 7% respectively) *Ibid.*, 270.

⁴⁹ Christine Liu. "Becoming a Christian Consciously Versus Nonconsciously," in *Journal of Psychology and Theology*, Vol. 19 No.4, (1991), 365. In 1990, using a "*Decision to be a Christian Questionnaire*," Liu polled 107 undergraduates enrolled at Stanford University, and found 26 (25%) had what she called the "non-conscious" decision, 69 (67%) the gradual, and eight (8%) the sudden decision. See *Ibid.*, 368.

gradual and those whose conscious decisions were sudden.⁵⁰

This literature review of the psychological literature revealed that there were basically two levels of conversion experience (sudden versus gradual) with two categories for each level. The Sudden grouping had Definite Crisis and Emotional Stimulus categories, whereas the Gradual grouping had Unconscious and Gradual Awakening categories.

But there is a third category of conversion experience that had not yet been clarified that is helpful in providing further insight into the gradual process. This category honors the theories of human development that embrace the concept of various stages of faiths that concur with the natural cognitive development in humans.⁵¹ It allows for the possibility of numerous commitments to religious beliefs throughout various stages in one's life. "Conversion can be sudden, but mostly it is gradual. Furthermore the validity of conversion is not found in the speed of the turning. It is found in the nature of commitment (to Jesus by repentance and faith)."⁵² Many will choose this option when given the opportunity to self-identify their conversion experience as part of a gradual process of committing one's life over and over again to Christ. Thus, the result is that five distinct categories on two different levels have been promulgated in attempting to empirically and operationally define Christian conversion

This brief but detailed review of the literature was provided to illustrate how extensively empirical methods are used to study one of the most important aspects of Christian spirituality – the conversion experience. But these quantitative studies do little to explain the complexity or diversity of the Spirit's move in this regard. The complex nature of human personality and the

⁵⁰ Ibid.

⁵¹ See James W. Fowler, *Stages of Faith: The Psychology of Human Development and the Quest for Meaning*. Reprint, 1981 (San Francisco: Harper, 1995) and James E. Loder, *The Transforming Moment* (Colorado Springs: Helmers and Howard Publishers, Inc., 1989) for excellent treatments of this subject.

⁵² Richard V. Peace, *Conversion in the New Testament* (Grand Rapids, Eerdmans, 1999), 302. Peace offers one of the most extensive treatments of the subject of "process evangelism" available.

various matrices of culture, family, and the Spirit's supervision of the interpersonal interaction between witness and respondent make attempts to quantify the Spirit's work seem overly simplistic. To reduce the whole of Christian conversion to just 5 categories, to a certain extent, denies the miraculous and mysterious nature of conversion. In this regard, attention to the individual narrative of conversion, as is addressed in qualitative analysis, is likely to be just as, if not more valid than larger studies that attempt to quantify the Spirit's activity in the heart of believers.⁵³

Quantitative and qualitative approaches to measuring the mystery of conversion in the hearts of humans inherently have limitations. The limitations in empirically defining such complex interactions of the Spirit as conversion play out in all kinds of confusion in the local church or missional context. Imagine the typical evangelist or his follow up team, after an altar call, polling people about what kind of conversion experience they had just gone through.

The Bad

Moving from the micro perspective of individual conversion to the macro of church-wide practices we see that not only are there problems with sampling biases and operational definitions of variables under study, but the reporting of statistics is commonly abused across Christianity to support the validity of ministries, outreaches, and revivals.

The so-called "Pensacola Portal" or revival at Brownsville Assembly of God in Pensacola, FL provides a case in point. This researcher was familiar with 3 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

⁵³ For one of the more comprehensive treatments of the complexity of conversion from various perspectives see Lewis R. Rambo, *Understanding Religious Conversion* (New Haven: Yale University Press, 1993).

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 that he had heard that over 300,000 people had come to Christ there since the revival's origin.⁵⁴ The widespread disparity in numbers converted is likely due not only to how conversion was operationally defined but who was polled concerning it.

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 raised a question by someone in the audience, "what happened to the other 113,500 believers whose conversion they had recorded?"

Unfortunately this case in point is probably more the rule than the anomaly when it comes to reporting conversions in larger crusade venues. The results are often based on anecdotal opinion rather than actual data gathering. Furthermore, the lack of common understanding of what actually constitutes a genuine conversion exacerbates the problem.

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. Many evangelists are not precise enough about what people are being asked to respond to in coming forward. The result is questionable data regarding the number of genuine first-time conversions. One of the more striking comparisons is the large number of conversions recorded over the last decade in the

⁵⁴ 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.

Assemblies of God according to the pastor-reported Annual Church Ministries Report, with little or no actual growth in church attendance as a result of all these conversions.⁵⁵

The Ugly

Another common myth heard 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. On the other hand, time and time again we hear that so many thousands or even millions have "Come to Christ" through someone's ministry. But what does that phrase actually mean?

Perhaps it means that they came forward in response to a call to have someone pray for them in general, perhaps for healing, or an unsaved loved one, or for some other pressing issue in their life rather than a need for conversion? In addition, there is considerable debate over the efficacy of "sinners' prayers" as a marker for genuine conversion. How many times have people prayed loudly with big crocodile tears of repentance at the conclusion of a church service – only never to be seen again by that congregation?

The church has traditionally seen the fruit of a consistent sanctified life as evidence of genuine conversion, more than passing responses to passionate altar pleas.⁵⁶ Leaders in the church and other missional contexts would do well to weigh through how they will operationally define genuine conversion if they really want to see long-term fruit in their ministries.

Once a clear and measurable definition of the variable(s) to be investigated has been completed, there are additional methodological issues that must be addressed to assure the

⁵⁵ 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, Assembly of God General Council, Springfield, Mo.

⁵⁶ The biblical accounts and church history reveal that baptism was one of the primary acts of obedience expected of new converts but in many Pentecostal and/or Evangelical churches passionate displays in response to altar calls as a form of public confession of sins seems to have taken its place.

veracity of any study of the Spirit's move in the heart of persons. Each of these issues are important in their own right as far as validity of findings are concerned and they will be addressed in the order that they are normally considered when developing experimental design.

Qualitative vs. Quantitative Analysis

The Good

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 into the mix.⁵⁷

Often times a combination of both quantitative and qualitative approaches are desirable to most accurately measure the variables in question.⁵⁸ For instance, recently a major research project funded by the Templeton Foundation investigating the effects of "divine love" on sacrificial service to others and its effect on revitalization within the Assemblies of God.⁵⁹ 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

⁵⁷ Gorsuch argues against the development of new measures when existing ones are more than adequate, see Richard Gorsuch, "Measurement: The Boon and The Bane of Investigating Religion," *American Psychologist*, 39, no.3 (1984), 234.

⁵⁸ 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." See Michael Q. Patton, *Qualitative Research & Evaluation Methods*, 3rd. ed. (Thousand Oaks, CA: Sage Publications, 2002), 585.

⁵⁹ "Margaret Poloma on Pentecostalism, the Assemblies of God, and Godly Love" from Research on Religion podcast, Dec. 6th, 2010, cf. <http://www.researchonreligion.org/protestantism/margaret-poloma-on-pentecostalism-the-assemblies-of-god-and-godly-love> (accessed January 15, 2012).

divine and human actors”⁶⁰ and it’s outplay within Pentecostal churches and organizations under study.

Other examples of the qualitative approach include participant observation, factor analysis, in-depth interviews, and content analysis. All of these can be helpful ways of assessing in more detail the move of the Spirit upon each individual on a case by case basis rather than gathering data from a broad spectrum of respondents and then attempting to deductively apply these to individual situations.

The Bad

There exists widespread use of both descriptive and inferential statistics from broad sampling frames that attempt to survey a nationwide population. Yet, these sampling frames may be mostly irrelevant to the issues that unique and anomalous individuals are facing on a day to day basis. Perhaps in these cases qualitative individual and longitudinal studies are much more applicable than broad generalizations extrapolated from scientific random samples. As Morgan and Morgan’s research shows “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.”⁶¹

In issues of individual conversion and/or spiritual development, qualitative approaches are likely to yield much more valid results than deductions drawn from the general population.

⁶⁰ Interview with Dr. Peter Althouse on June 12th, Southeastern University, Lakeland, FL. Althouse is one of the researchers involved in the study with Paloma.

⁶¹ David L. and Robin K. Morgan, “Single Participant Research Design: Bringing Science to Managed Care,” *American Psychologist*, (2001): 122.

What pastoral counselor would attempt to console a client experiencing deep psychological wounding with some generalized data from a national study concerning that issue?

The Ugly

All of us have likely heard at least one preacher or even a teacher in the classroom who injudiciously uses an anecdotal case of one to extrapolate to the broader congregation as a whole. Similarly, we will commonly read of research that started out as a qualitative study, with just 2 or 3 participants, a technique called triangulation,⁶² but then in conclusion attempts to extrapolate those findings to the more general population. All of us have a tendency to want to read generalizations back into our own situations. Yet good research militates against this tendency by warning the readers of the limitations of their study and the methodology employed.

Descriptive vs. Inferential Statistics

The Good

It may be helpful to understand some common nomenclature used by statisticians to understand some of the basic issues in choices between the two basic types of statistical analysis: Descriptive and Inferential.⁶³

Descriptive statistics are for organizing and summarizing sample data so we can *describe* the important characteristics concerning them. Common descriptive statistics are the number in the population being measured (N), the number in the subset or sample (n), the percentile rank or percentage that number represents (%), the Mean (M) or average of all the numbers in the sample, the median (Mdn) or point at which 50% of the scores in the sample are above or below, and the mode (Mde), the most frequent number in the sample.

⁶² Patton reports there are at least four basic types of triangulation: 1) data triangulation, 2) investigator triangulation, 3) theory triangulation, and 4) methodological triangulation, Patton 247.

⁶³ For one of the better treatments of these issues see Gary Heiman, *Basic Statistics for the Social Sciences* (Independence, KY: Cengage Learning, 2010).

Beyond descriptive stats are more sophisticated statistical procedures involving the use of inferential statistics. These are used to make *inferences* about scores and/or relationships represented by the sample. One of the most fundamental inferential statistics is derived from a mathematical calculation of relationships or associations between cases or individuals in a population where no manipulation or intervention is involved. It is known as the correlation coefficient (r). Comparisons between the average groups of cases in a sample are either the t-test (t) when only 2 groups are involved or the F-test (F) or Analysis of Variance (*ANOVA*) which involves a comparison of means between 3 or more groups.

Inferential stats are calculated on the basis of what is known as Null Hypothesis Significance Testing (*NHST*). This is a basic calculation that determines if the results are due to error or chance and presented in terms of the probability that such a result is due to error. The basic standard is that the probability (p) that the result is due to error or chance should be less than 5%, which is known as a 95% confidence interval and represented as $p < .05$.

The Bad

Basic statistics lessons aside, there are a number of errors that commonly surface when dealing with statistical calculations. A common one is assuming that there is a *statistically significant* finding when there is none.⁶⁴ For instance, you might assume that if the mean or average of the high school GPA's of incoming males ($n = 5$) in a university was 2.75 and the females ($n = 7$) was 3.25, that this large difference would be significant. However, you must look closer because the sample size may have been so small that the result would not be statistically significant.⁶⁵ Yet without a clear understanding of this critical issue the assumption of a

⁶⁴ The phrase is a technical one used by statisticians and is true only if $p < .05$.

⁶⁵ Generally at least 35 subjects ($n = 35$) are necessary for a result to be statistically significant. This is especially true for the Pearson r correlation coefficient that is one of the more "robust" inferential procedures. Generally for

meaningful difference may go unchallenged.

Another common misconception is in the use of correlations, where you might have a statistically significant result, where $p < .05$, but the relationship is so weak as to be negligible. The same can be true when sample sizes are very large on mean comparisons, the probability the result is due to error could fit the standards of NHST but be essentially meaningless. Using the previous example, you could have male GPA's with a $M = 2.99$ and female GPA of 3.00 and if the sample size is large enough the likelihood of statistical error could be ruled out and it would be considered statistically significant but practically miniscule.

The Ugly

Two examples can be given of worst case scenarios in the use of descriptive vs. inferential statistics. The first is where no attempt is made to test the statistical significance at all but it is assumed anyway. The use of simple descriptive stats can result in an implied sophistication even though no probability level is provided or substantial inferential test completed.

Another example is where statistical analysis has been completed but is indecipherable to the common person and even those with basic knowledge of empirical science. Some studies are so in-depth and complicated and the results presented in such a complicated manner that only someone with a Ph.D. in statistics, which is a language all its own, could possibly make sense of it. Anyone who has read any number of scholarly journal articles in the social science genre has encountered this kind of reporting. The research could be valid and helpful, but so technical in its reporting that a very minute segment of the population can even interpret the results. Such presentations are often indicative of a "paralysis of analysis," - reporting of minutiae that make

means comparisons; t and F tests, at least 35 in each group would be the minimum desirable for a result to yield statistical significance although there are many other factors that influence this as well.

little if any difference in the real world because so few can access the data for practical application.

Sampling Frames

The Good

Once one has accurately and operationally defined the variables and decided on quantitative and/or qualitative approaches, then a sampling frame needs to be selected. How this is done can greatly influence the outcome. Pollsters rely on the null hypothesis significance theory in statistics to arrive at 95% confidence intervals. The idea is that if you conduct a “scientific” or “simple random” sample, there is a 95% likelihood that every person in that population (the U.S.A. let’s say) 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 Bad

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 that are not church attendees or believers and extrapolated those findings to the entire U.S. population.⁶⁶

⁶⁶ Thom S. Rainer, *The Unchurched Next Door: Understanding Faith Stages as Keys to Sharing Your Faith* (Grand Rapids, Mich.: Zondervan, 2003). In Rainer’s defense he argues that his sample “represented every conceivable demographic area of America,” *Ibid.*, 18. His methodology was supported by Professor Jon Rainbow a research and statistic consultant, who acknowledged the approach was qualitative in design. In the Appendix B: Research Design

Choices regarding who to poll can have dramatic impact on responses because of the nature of the respondents. For instance, when polling un-churched persons about their attitudes regarding the church in general some stark differences in opinions emerge just based on 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 that were not likely to ever be a part of a church.⁶⁷

Rainer, for research in *Surprising Insight From the Formerly Un-churched* had a much different approach.⁶⁸ He polled 353 people that “for all or a large portion of their lives were not in church” but now are Christians attending church.⁶⁹ He surveyed those who are now a part of a local church on why they became so. The construct validity of this experimental design is much more reliable. It addresses some of the same issues as Kimball, but draws data from a much more telling sample.

Kinneman, 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.⁷⁰ 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.

and Statistical Review it states emphatically “no generalization is attempted or implied,” Ibid., 252. Yet, the book clearly extrapolates the findings to the U.S. population at large throughout.

⁶⁷ Dan Kimball, *They Like Jesus but Not the Church: Insights from Emerging Generations* (Grand Rapids, Mich.: Zondervan, 2007).

⁶⁸ Thom S. Rainer, *Surprising Insights from the Unchurched and Proven Ways to Reach Them* (Grand Rapids, Mich.: Zondervan, 2001).

⁶⁹ Ibid., 22.

⁷⁰ David Kinneman and Gabe Lyons, *Un-Christian: What a New Generation Really Thinks About Christianity and Why It Matters*, (Grand Rapids: Baker Books, 2007), cf., 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 $n = 613 - n = 1020$) and self-select online polls (from $n = 102$ to $n = 2409$).

He found that 84% of his respondents, who were young people between the ages of 13-29 years, know a Christian, but only 15% see a difference.⁷¹ Perhaps this is indicative of a significant design assumption – that how one defines a Christian is largely left to self-identification as such.

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.”⁷² 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.⁷³ Among the so-called self-identified born-again Christians polled for the *UnChristian* study 59% of born-again Busters (ages 23-41) think cohabitation is morally acceptable compared to 33% of born-again older adults (age 42+). Consistently, 44% of those same Busters and 23% of the older cohort see sex outside of marriage also as morally acceptable. Nearly a third (28%) of those Busters thought having a sexual relationship with someone of the same sex was morally acceptable, and even 13% of the older adults agreed. If

⁷¹ Ibid., 48.

⁷² George Barna, *Third Millennium Teens: Research on the Minds, Hearts and Souls of America's Teenagers* (Ventura, CA: Barna Research Group, 1999).

⁷³ 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. cf., Ibid. 247.

this data is indicative of all who self-identify as Born-Again Christians - no wonder the culture at large sees any difference!

The Ugly

One of the more frightening examples of an abuse of statistics related to respondent bias is the pervasive use of course evaluations drawn from either too small or too biased sampling frames to assess instructor proficiency. While a self-report measure may indeed offer a reliable snapshot of what happened throughout that course there are some real dangers inherent to this method.⁷⁴ Such self-reports have a self-select bias built into them when they are not compulsory for all students.⁷⁵

It's possible that the type of student who is most likely to complete course evaluations are overly conscientious and certain teachers will not do well when evaluated by those personalities. Others may respond out of vindictiveness because of poor performance on a class assignment, especially if the evaluations are done during finals week. Furthermore, there can be many confounds that might influence how respondents rate the course, e.g., the time of day the class meets, the number of people in the class, the nature of the cohort, etc. All of these extraneous variables are related to challenges stemming from a skewed sampling frame. Many of these are beyond the control of the instructor. However, in some institutions student course evaluations are the primary means of assessing instructor effectiveness for promotion or even retention, yet remain beyond the authority of the faculty being assessed.

⁷⁴ For a thorough analysis of issues pertaining to self-report bias see Norbert Schwarz, "Self-Reports: How the Questions Shape the Answers" *American Psychologist*, (1999): 93-105.

⁷⁵ It could be argued that forcing students to complete the evaluation will also skew the results in a negative manner.

Subjective Self-Reporting

The Good

Self-report measures, such as questionnaires and surveys may be the most effective way to assess certain subjective experiences such as conversion, Spirit-baptism, sense of divine love, and other existential moves of the Spirit. The ease of access to large sample sizes through online and face to face polling, coupled with computer assisted statistical analysis, makes quantitative methods more efficient than ever. For whatever the reason, self-report instruments in the assessment of spiritual experience are abundant in religious research reporting, but each and every research project needs to be critically assessed to avoid the pitfalls inherently a part of this methodology.

Advantages of Self-Report Measures

While focus group interviews are often sometimes employed to assess spirituality⁷⁶ 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 presently 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 upon declaration of the major, which is compared later to the same post-test assessment upon graduation.⁷⁷

Perhaps one of the more popular self-report measures is the questionnaire compiled by

⁷⁶Cartledge uses this technique of recording group interviews then mining the transcriptions for insightful data, see Mark Cartledge, *Testimony in the Spirit: Re-scripting Ordinary Pentecostal Theology* (Burlington, VT: Ashgate Publishing, 2010).

⁷⁷ 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, for an exploratory and confirmatory factor analysis of these domains with other common psycho-metrics.

Schwarz.⁷⁸ His organization, Natural Church Development, gathered data from over 70,000 churches in more than 83 countries of the world on all six continents, 35,000 of these churches in the U.S. alone.⁷⁹ The senior pastor and 30 key lay people he or she selects from the congregation are polled. Data analysis provides churches with not only an internal comparison of the Eight Quality Characteristics of church health, but also offers comparisons between growing, healthy, unhealthy, and not growing churches on these factors.⁸⁰

Data from these surveys provides the local church with a composite “snapshot” of corporate spirituality by comparing and helping church leaders to identify the minimum factor from among these characteristics in their congregation in order to focus improvement on that factor. The results also provide empirical data for the local church leaders based on comparisons with the data for other congregations within their denomination, region or country.

While such self-report measures may be informative they are only helpful if the church takes the data and actually does something with it. One of the critiques of the NCD survey is that while thousands of churches have completed the survey, a relatively small percentage have actually gone through a strategic change cycle - which is what is recommended once the minimum factor has been identified.⁸¹

The Bad

Faking Good

Social scientists are familiar with a common response intended to subconsciously please

⁷⁸ Christian A. Schwarz, *Color Your World With Natural Church Development: Experiencing All that God Has Designed You to Be* (St Charles, Ill: ChurchSmart Resources, 2005).

⁷⁹ See <http://www.ncd-international.org/public/essence.html>, (accessed January 17, 2012).

⁸⁰ 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, *Ibid*. The emphasis is put on the adjective as the primary operational definition, Schwarz, 105.

⁸¹ NCD claims that “churches that have completed three or more NCD surveys have increased their average growth rate by 51% between the first and third survey, <http://www.ncd-international.org/public/essence.html>, (accessed January 17, 2012).

the interviewer on self-report measures known as “faking good.” As a classic example of this phenomenon, Americans overwhelmingly identify themselves as Christians but apparently do not practice what they preach as far as church attendance is concerned.⁸²

Rainer reports that over 80% of those who were not Antagonistic or Resistant to Christianity in one of his studies would be interested in a Bible study or learning more about Jesus. Furthermore, his data purports that over 80% of all his respondents including those who were Antagonistic or Resistant “said they would come [to church] if invited by a Christian.”⁸³ However, responding this way to a researcher who asks such things over the phone is very different than being asked more directly for a response that is measurable by actual behavior rather than intent, i.e., “would you like to attend a Bible study in my home or attend my church with me this week?” The reader may intuitively sense there is a decided disconnect between espoused belief and actual practice in the American religious psyche.

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% of Americans attended church on Sunday, contend that the number is likely only half that because of the way the data is gathered – through self-report measures.⁸⁴ For instance, when researchers polled a congregation in Ohio about their Sunday School attendance for the prior Sunday, 60% said they had attended when only 38% had actually done so.⁸⁵ Pastors and church leaders, who are often promised participation by their congregants

⁸² 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 January 17, 2012).

⁸³ Rainer, *Unchurched Next Door*, 232.

⁸⁴ See C. Kirk Hidaway, 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 Hidaway, 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.

⁸⁵ Hidaway, Marler and Chaves, 127.

at various events or programs, only to find that a much smaller number of volunteers actually show up, are not likely to be too 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 insightful and comprehensive studies available concerning self-report measures found that question content, format, and ordering conditions have a profound impact on responses.⁸⁶

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% rated themselves between -5 and 0, but only 13% endorsed the formally equivalent score on a scale from 0 to 10.⁸⁷

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 “source of information that respondents tend to draw on in order to determine their task and to arrive at a useful and informative answer.”⁸⁸ In other words, due to the respondents’ desire to do well, the way the questions are asked, the potential answers that are offered, even the order of the questions can have a profound impact on the results.

The Ugly

One of the more common pitfalls seen in the use of self-report surveys refers back to the issue of sampling frames. Many self-report respondents are not part of a scientific random

⁸⁶ Schwarz, “Self-Reports.”

⁸⁷ Ibid., 95.

⁸⁸ Ibid., 103.

sample which would give everyone in the U.S. a 95% chance of being sampled. Rather, what is more common are measures based on a haphazard or convenience sampling which are not truly representative.

For instance, when gathering data in a study of the relationship between conversion beliefs and evangelism practices this researcher was able to obtain answers from some 845 respondents.⁸⁹ However, this was not a representative sample of Christians across the U.S. as the instrument was administered by pastors in churches sympathetic to the researcher.

However, regardless of caveats to the contrary concerning the external validity of the research, many readers will extrapolate these findings, which are valid for this limited sample, to the greater population of the whole. Such an uncritical wholesale application to broader populations from small unrepresentative samples is all too common.

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. I have tried to provide here a primer of sorts for the novice in research methods to help in their assessment of empirical approaches to how the Spirit is moving in the lives of God's people.

I have tried to show that critical thinkers should carefully analyze 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 couple that with a wariness to accept every assertion that is made through the use of statistical data. By doing so, religious researchers will be able to avoid many of the pitfalls inherent to the use of empirical methods in studying the Spirit's activity.

⁸⁹ For a brief discussion of these results see Brian M. Kelly, "Practical Theology and Evangelism," *Journal of the Academy for Evangelism in Theological Education*, (1998-1999), 21-22.

The Lord Jesus admonishes us to bear fruit that will last (Mt.7:15-20). This statement infers some form of measurement and assessment of the everlasting fruit he has chosen us to yield. However, the use of empirical methods is inherently limited when it comes to an investigation of the Spirit's guidance in everyday Christian practice. The existential aspects of these interactions necessarily limit finite human abilities to measure the Spirit's movements, such as; the nature of numinous encounters, spiritual development in persons, the complexity of the conversion process, validity of divine healing experiences, the role of prayer, and the nature of the Spirit's interaction in cognitive, behavioral, and emotive realms. However, in spite of these limitations we should not shrink back altogether from trying to measure and analyze what God is doing in our world and in our churches. Rather, by empowering the body of Christ to critically engage its use of empirical methods we can have confidence that the fruit we are claiming is indeed fruit that will last.