Annual Meeting of the *Rapaport-Klein Study Group*
Austen Riggs Center, Stockbridge, Massachusetts, June 11-13, 2010

Panel discussion on "Clinical and Empirical Issues: Disagreements and Agreements"
(The background for this panel is the controversy stirred by Irwin Hoffman's article
"Doublethinking our way to 'scientific' legitimacy: The desiccation of human experience."
Journal of the American Psychoanalytic Association, 2009, 57, 5: 1043-1069)

Discussion by David L. Wolitzky

(Working copy, DRAFT B 3/19/10;5/12/10, NOT FOR CITATION)

1) Shorter version (*For purposes of our meeting, here at the beginning I am also posting a short critique using selected points from my longer critique of Hoffman's paper. It will be sufficient to read this short version in lieu of the longer paper previously posted. The material below was written for a Newsletter edited by Jeremy Safran. The longer version is after this shorter version)*

At the invitation of the Editor, I have selected and modified excerpts from a lengthy critique (in preparation) of Hoffman's (2009) paper, "Doublethinking Our Way to "Scientific Legitimacy": The Desiccation of Human Experience." Hoffman (2009) argues that to accord "privileged status" to "systematic research and neuroscience as compared with in-depth case studies… is unwarranted epistemologically and potentially damaging both to the development of our understanding of the analytic process itself and to the quality of our clinical work." (p. 1043). Hoffman's (2009) use of the Orwellian phrase "doublethink" in the main title of his paper reflects his view that scientific approaches to psychoanalytic theory and practice reflect a self deceptive capitulation to political pressures.

The subtitle of his paper---The Desiccation of Human Experience (emphasis added)---communicates his view that research on the psychoanalytic process denigrates the uniqueness and "limitless complexity" of the human being. Although Hoffman is careful to note that he is not claiming that research has "no" value (p. 1044, original italics), he appears to believe it has very little and that it is fundamentally anti-humanistic. The polemical, rhetorical nature of Hoffman's paper suggests that he regards virtually any use of numbers or categories in relation to patients' experiences as a "desiccation" of human experience and, *ipso facto*, a violation of the patient's dignity.

Empirical research and clinical case studies

Although history has repeatedly shown that any method or procedure (whether research or clinical) can be deliberately or unwittingly misused for nefarious purposes, (e.g., in the service of Orwellian, authoritarian thought control), the advantage of science with its emphasis on accessibility to observable data, replication, and controlled conditions is, in principle, a useful safeguard against being ruled by dogma and blind obedience to a persuasive, charismatic leader's point of view. Yet, as Hoffman sees it, scientific studies of psychoanalysis put us on the road to "authoritarian objectivism" and to a "conformist" rather than a "critical" psychoanalysis.

Politically inspired "scientism" unfortunately can be used tendentiously in the service of seeking premature closure as to which treatments are effective for which conditions. However there are a number of error-correcting procedures that serve as 'checks-and-balances' against rampant "scientism". For example, analytic researchers are well aware of the distinctions between statistical and clinical significance, between the efficacy and effectiveness of treatment, and of the limitations of randomized controlled trials. They are not guilty of engaging in the "doublethinking" Hoffman attributes to them; as indicated, they clearly are aware of both the value and limitations of empirical research whereas in true doublethink the person holds contradictory beliefs without awareness of the contradiction.

Furthermore, the shift to the idea of "evidence-based practice" refers to the value of integrating research evidence **and** clinical expertise **and** patient values. Adopting this approach should not compromise the quality of clinical work, as Hoffman fears, as it would seem to leave ample room for the "art" of psychoanalysis. From this perspective, case studies and empirical research can be considered as complementary methods of advancing psychoanalytic knowledge. Each method has something unique to offer and each method has its limitations, depending on the kind of question being asked. Thus, there are some important questions for which research clearly deserves to have privileged status over clinical case reports, and even over the systematic study of a few single cases. To cite but two examples: (1) the outcome of analytic treatment at termination and at follow-up intervals, and, (2) the kinds of patient and therapist variables associated with differential treatment outcomes cannot be answered by exclusive reliance on clinical case studies.

Given that Hoffman wished to extol the virtues and epistemological sufficiency, even superiority of the clinical case study method compared with systematic empirical research, one would have hoped that he would have argued his position more effectively by offering a nuanced critique of the flaws in the case study method and a discussion of the means available for enabling case studies to have greater probative value. For example, Kachele, Schachter, and Thoma (2009) present a compelling argument for the systematic study of single cases and their advantages over typical clinical case reports.

Accountability

To suggest that research on therapy outcome is merely a capitulation to political pressure overlooks the point that even if the wider society did not demand accountability, our own sense of morality requires that we back up our assertions with evidence that goes beyond clinical case reports. Strong, or even modest, claims of therapeutic effectiveness based only on clinical case reports can readily be dismissed as analysts' self-congratulatory testimonials.

The shift in conceptualization of the psychoanalytic situation from a one-person psychology aimed at interpretations that "tally with what is real" to a two-person, hermeneutic, constructivist view of the analytic situation in which the coherence of the co-constructed narrative is primary, does not absolve analysts from satisfying the need for accountability. Even advocates of a constructivist view, who contrast their view with what they denigrate as "objectivism", are making at least implicit claims that the treatment they offer is effective and probably more effective than treatments inspired by other perspectives. How do they know? How *can* they know without supplementing their clinical judgments with independent assessments of therapeutic benefit and the stability of therapeutic gains? And, how can Hoffman, or any individual practitioner, confidently claim that privileging empirical research will decrease the quality of clinical work without someone conducting research on the current quality of clinical work and tracking whether it is declining over time, as Hoffman fears it will?

Concluding comment

Hoffman failed to even consider the kinds of question for which empirical research might deserve to have privileged epistemological status over clinical case reports. Nor does he grapple with the methodological problems that beset case studies and compromises their evidential value. The result is a markedly tilted view of how reliable psychoanalytic knowledge is to be gained.

One can appreciate that his claims probably strike an emotionally resonant chord with many colleagues. However, to accuse those who engage in or respect psychoanalytic research of “authoritarian objectivism”, or of being intolerant of ambiguity, is to overlook the comparable dangers of exclusive reliance on whatever clinical/theoretical thinking is in vogue. To cite but three examples, consider the excessive human suffering caused by clinicians tenaciously clinging to the view that homosexuality is pathological, that schizophrenia is caused by bad mothering (i.e., the “schizophrenogenic mother”), or that autism is the result of a “refrigerator mother”. In each case, it took systematic, methodologically sound, empirical research (aided by societal pressures) to dislodge the clinically based views.

The point is that both systematic empirical research and clinical case studies can contribute to our understanding of the process and outcome of psychoanalytic treatment as well as to psychoanalytic theories. But, to pit the two methods against each other is not constructive as each is designed to make different kinds of knowledge claims. For instance, if we want to make general statements about the outcome of treatment and the stability of therapeutic gains, we need to privilege research, as clinical experience alone cannot provide such answers. If we are seeking an in-depth understanding of an individual patient, the multiple meanings of his or her conflicts, dreams and fantasies, then the case study method is more suitable. But, to protest that research findings are of no immediate help to the therapist and to engender fear that research will decrease the quality of clinical work is to sound false alarms that are detrimental to psychoanalysis. Actually, most clinicians do not even read the research literature, and if they do, they tend to dismiss it, particularly if they encounter findings contrary to their cherished beliefs. Of course, when called upon to defend their work, therapists are happy to cite empirical research that demonstrates the effectiveness of dynamic psychotherapy (e.g., Shedler, 2010).

Gill (1994, p. 157) observed: “We may be satisfied that our field is advancing, but psychoanalysis is the only significant branch of human knowledge and therapy that refuses to conform to the demand of Western civilization for some kind of systematic demonstration of its contentions”. In the past decade, analytic researchers increasingly have responded to Gill’s challenge, with no apparent indications that clinical work has suffered as a result, an outcome feared by Hoffman. There is no inherent incompatibility in using clinical and research methods to enhance our understanding, if both are conducted in a sophisticated, disciplined manner, addressed to questions they are best suited to answer, and “privileged” accordingly. Hoffman’s concerns about “scientism”, though overstated, are mainly valuable for clinicians whose grasp of research methodology is tenuous or has eroded from years of disuse.

In conclusion, the future of psychoanalysis is not well served by pitting research against clinical case studies in such a polarized manner and associating the former with a host of related vices such as “authoritarian objectivism”, doublethink, endangering the quality of clinical work, and so on, in contrast to the virtues of free will and humanistic values that uphold and cherish the dignity and uniqueness of the individual.

References

Gill M.M. (1994). *Psychoanalysis in Transition: A Personal View*. Hillsdale, NJ: Analytic Press.

Hoffman I.Z. (2009). Doublethinking our way to “scientific” legitimacy: The desiccation of human experience." *Journal of the American Psychoanalytic Association*, 57, 5: 1043-1069. Internet edition: http://www.psychomedia.it/rapaport-klein/hoffman-09_Doublethink.pdf.

Kachele H., Schachter J. & and Thoma H. (2009). *From Psychoanalytic Narrative to Empirical Single Case Research: Implications for Psychoanalytic Practice*. New York: Routledge.

Shedler J. (2010). The efficacy of psychodynamic psychotherapy. *American Psychologist*, 65, 2: 98-109. Internet edition: http://www.apsa.org/portals/1/docs/news/JonathanShedlerStudy20100202.pdf.

2) Longer version

Introduction

Hoffman's (2009) lead paper in this journal was based on his plenary address to the *American Psychoanalytic Association* in 2007, a talk that was enthusiastically endorsed by the audience as evidenced by the standing ovation he received. The overarching theme of his paper was to warn of the dangers that would ensue if systematic empirical research were to be privileged over the clinical case study method as the preferred path to psychoanalytic knowledge.

Although we believe that the concerns raised by Hoffman are not entirely without merit, we feel he has grossly overstated them. Because we believe his views are largely misguided and detrimental to the future of psychoanalysis, yet apparently wholeheartedly embraced by many colleagues, we thought it important to offer a detailed rebuttal of Hoffman's position as well as to present more general comments concerning the issues raised by him. For those with a strong interest in these issues it might prove helpful to read the remarks to follow in conjunction with a review of Hoffman's (2009) article.[1]

We do not object to Hoffman's obvious passion in expressing his beliefs. However, it detracts significantly from the substance of his argument when his passion leads him to rhetorical excesses in an effort to persuade readers of his views. His indulgence in rhetoric begins with the title and subtitle of his paper--- *Doublethinking* Our Way to "Scientific Legitimacy": The *Desiccation* of Human Experience (italics added). The phrase "doublethinking" is a reference to Orwell's novel, *1984*, in which social thought control becomes the order of the day. The phrase "desiccation of human experience" implies that to measure and to quantify aspects of the psychoanalytic encounter immediately and inevitably destroys its purity and uniqueness as well as the dignity of the individual. In our view, Hoffman's (2009) paper is more a political tract and passionate polemic advocating exclusive reliance on the clinical case study method than a careful consideration of the relative merits and limitations of clinical case reports versus systematic empirical research.

We should note at the outset that we share Hoffman's humanistic views and can appreciate that they inspired his paper. We also feel that Hoffman is clearly correct in alerting us to the danger that science can be perverted into "scientism" in pursuit of undue ideological power and influence. At the same time, it seems to us that Hoffman is so alarmed by this potential danger that he places psychoanalysis practically off limits for scientific study and, as noted above, fails to offer a nuanced account of the virtues, limitations, and dangers of both systematic empirical research *and* the clinical case study method, nor does he suggest a synergistic, complementary role for both approaches.

We will organize our discussion around the key points raised by Hoffman but also will offer some general comments about the virtues and limitations of case studies and of empirical research. We would have thought that these remarks would be unnecessary and non-controversial at this point in psychoanalytic history, but the apparently widespread, uncritical acceptance of Hoffman's polarized view of 'empirical research' *versus* 'clinical case studies' views makes our comments a hopeful corrective.

---

[1] Of course, Hoffman is far from alone in being vehemently opposed to psychoanalytic research. For instance, Green (2001, p. 21) writes "the noble term 'research' carries such an amount of prestige that it is to be expected that any reference to it might compel one to bow before it. Unfortunately, compared with the richness of the clinical experience of psychoanalysis, the findings of researchers look very meager".

The privileged status of case studies versus empirical research

Hoffman (2009) argues that there is a danger that “privileged status” will be accorded to “systematic research and neuroscience as compared with in-depth case studies and strictly psychological accounts of the psychoanalytic process”. He claims that such privileged status “is unwarranted epistemologically and potentially damaging both to the development of our understanding of the analytic process itself and to the quality of our clinical work.” Hoffman’s (2009) use of the Orwellian phrase “doublethink” in the main title of his paper seems to reflect his view that scientific approaches to psychoanalytic theory and practice reflect a self deceptive capitulation to political pressures. He juxtaposes quotes from Orwell’s *1984* definition of “doublethink” with Fonagy’s (xxxx) urging that we take seriously the need for accountability and therefore for research.

Although history has repeatedly shown that any method or procedure (whether research or clinical) can be deliberately or unwittingly misused for nefarious purposes, (e.g., in the service of Orwellian, authoritarian thought control), the advantage of science with its emphasis on accessibility to observable data, replication, and controlled conditions is, in principle, a better safeguard against being ruled by dogma and blind obedience to the “leader’s” point of view. As one example, the finding of an “allegiance effect” in therapy outcome studies (i.e., the outcomes are more apt to accord with the investigators’ theoretical orientation [Luborksy *et al*., 1999, 2002]) allows us to control for this factor in subsequent studies. It is much more difficult to deal with the “allegiance effect” in clinical case studies.

There is a valid concern that politically inspired “scientism” can be used tendentiously in the service of seeking premature closure, as in the early “empirically validated” or “empirically supported treatment” (EST) movement, a movement biased in favor of certain therapies. For example, to coerce training programs to train students only in therapies that have so far received support in ‘efficacy’ studies would be a miscarriage of science for self-serving, ideological, competitive purposes. It would “freeze” whatever therapeutic procedures are in favor at a given point in time and preclude innovations and improvements.

However, there are a number of error-correcting procedures that serve as ‘checks-and-balances’ against rampant “scientism”. First, we now speak of “evidence-based practice”. “Evidence-based practice” is the “integration of best research evidence with clinical expertise and patient values.” Adopting this approach should not compromise the quality of clinical work, as Hoffman fears. In medicine, for example, there is no inherent incompatibility between researchers learning more about the relative effectiveness of different medications (and dosages) for hypertension and physicians, aware of this research, deciding which drug or combination of drugs at what dosage levels appear most optimal for a given patient given that patient’s unique medical conditions and history. Thus, although diuretics might be the first drug to be considered, there are circumstances under which that could be a bad choice. Of course, it helps if those circumstances have some research basis, but the point is that there is still a significant role for clinical judgment. It also is understandable that clinicians can conclude that certain research findings are not relevant to their patient. For example, although studies show a statistically significant relationship between the quality of therapeutic alliance and treatment outcome, that relationship is modest. However, the analyst is free to feel that in some, if not all, of his or her cases, the state of the alliance is not only an important factor but a critical one. Unless Hoffman is afraid of an Orwellian Big Brother monitoring of what the analyst does in the privacy of his or her office or that strict adherence to a specific treatment manual will become the order of the day, it is not clear why we cannot value research findings *and* clinical expertise, along with patient values. This view would certainly seem to leave ample room for the “art” of psychoanalysis.

From this perspective, case studies and empirical research can be considered as complementary. Each method has something to offer and each method has limitations, depending on the kind of question being asked. Yet, Hoffman fears that, although the clinical case study method has been the dominant basis of how clinicians approach their work as therapists, the research approach will overshadow the clinical case study, be accorded "privileged" status, and lead us astray.

We believe there is ample protection against the misuse of research data and that analytic researchers decidedly are not engaged in "double thinking", as Hoffman alleges. In true "double think", the person is not even aware of holding contradictory views. With regard to analytic research, investigators show a clear awareness of the contributions as well as the misuses of their work. For example, researchers are well aware of the distinction between "efficacy" and "effectiveness", noting that results obtained in randomized controlled trials, following a strict treatment protocol, with patients without comorbid conditions yields results that are not readily generalizable to everyday clinical practice in the real world. They also offer cogent criticisms of the EST movement (e.g., Westen, Morrison Novotny & Thompson-Brenner, 2004). Analytic researchers also readily endorse Kazdin's (xxxx) persuasive point that that we need to avoid "arbitrary metrics", i.e., we need to distinguish between statistically significant changes and clinically meaningful changes.

Thus, there is a growing cadre of sophisticated psychoanalytic therapy researchers whose work is far from simplistic. These research clinicians are doing us a service by showing that our treatment results are at least as good as those achieved by medication or by CBT (e.g., Shedler, 2010). If we also can offer a solid body of data concerning superior post-treatment consolidation of treatment gains and greater stability of such gains as a result of psychoanalytic treatment, we will have met the challenge of accountability and we can easily justify being interested in more than symptom amelioration. It is this kind of work that could win the day for psychoanalysis.

Accountability

To suggest that research on therapy outcome is merely a capitulation to political pressure overlooks the point that even if the wider society did not demand accountability, our own sense of morality requires that we back up our assertions. As Gill (1994, p. 157) observed, "We may be satisfied that our field is advancing, but psychoanalysis is the only significant branch of human knowledge and therapy that refuses to conform to the demand of Western civilization for some kind of systematic demonstration of its contentions".

Strong, or even modest, claims of therapeutic effectiveness that are based only on clinical case reports can readily be dismissed as analysts' self-congratulatory testimonials. It is true that CBT clinicians will even distrust and distort our empirical studies in an effort to retain their growing hegemony. But, bias expressed in scientific articles can be countered and can prevent science from being used to advance a self-serving agenda (e.g., the finding of an "allegiance effect" in outcome studies). The surest way to see analytic therapy disappear as a worthy contender in the therapy arena is not to pursue research on its effectiveness. Those who do conduct such research are not engaged in "doublethink", as Hoffman would have us believe. Doublethink entails "simultaneously accepting as correct two mutually contradictory beliefs" and not even realizing that one is doing so. This is hardly an accurate characterization of psychoanalytic researchers who are fully aware of the importance, the limitations, and the potential misuses of empirical research.

The shift in conceptualization of the psychoanalytic situation from a one-person psychology aimed at interpretations that "tally with what is real" to a two-person, hermeneutic, constructivist view of the analytic situation in which the coherence of the co-constructed narrative is primary, does not absolve analysts of satisfying the need for accountability. One can and should still search for lawful regularities concerning the process and outcomes of psychoanalysis. Even advocates of a constructivist view, who contrast their view with what they denigrate as “objectivism”, are making at least implicit claims that the treatment they offer is effective. How do they know? How can they know without supplementing their clinical judgments with independent assessments of therapeutic benefit and of the stability of therapeutic gains? Yet, if we were to follow Hoffman’s approach, such studies would never be done. Of course, when they are done by analytic researchers clinicians do not hesitate to cite them as evidence supporting their clinical work (e.g., Shedler, 2010).

The “desiccation” of human experience

The tone and content of Hoffman’s paper suggests that he regards virtually any use of numbers or categories in relation to patients constitutes a “desiccation” of human experience. For example, although the authors of the PDM clearly are aware of the limitations and inevitable oversimplification of any classification system, of the artifactual nature of the high co-morbidities in the DSM-IV-TR, and of the tendency to “reify complex syndromes” (PDM, p. 31n), Hoffman (2009, p. 1060) views the PDM approach as merely a “nod to humanistic, existential respect for the uniqueness and limitless complexity of any person” because, like the DSM, the PDM manual gives code numbers to the different diagnostic entities. Surely, psychoanalysts using the PDM know that, as with the DSM, the vast majority of patients do not meet the full diagnostic criteria for a single disorder but show characteristics of several disorders.

There are a number of statements in the introduction to the PDM that reflect the desire to create a clinically meaningful approach while acknowledging the difficulties of doing so. For example, the authors state (p. 5-6) that there is “a healthy tension between the goals of capturing the complexity of clinical phenomena (functional understanding) and developing criteria that can be reliably judged and employed in research (descriptive understanding)”. Although the authors believe “It is vital to embrace this tension…” (p. 6), if one does not share the second goal, and Hoffman does not, there is no need to “embrace this tension”.

Another example of Hoffman’s aversion to measurement is his critique of the Shedler-Westen Assessment Procedure (SWAP) (Shedler & Westen, 2006). The SWAP is an attempt to use clinical judgment in personality assessment in a manner that is “…both dynamically relevant and empirically grounded” (Shedler & Westen, 2006, p. 576). Apparently, Hoffman is not impressed that the SWAP is the result of a “7-year interative revision process” (p. 580) in which more than 2000 clinicians were sampled in an effort to find a set of personality descriptive items that clinicians of all theoretical persuasions felt captured important aspects of their patients’ personalities. The SWAP contains 200 items. A minority of clinicians (between 5% and 20%, depending on how one interprets the scale points used by Shedler & Westen [Shedler, personal communication, February 2010]) did not “agree” or “strongly agree” that the items captured what they felt was important. Hoffman said he would “put my money” on this minority. He believes that those therapists are likely to be the ones who are “profoundly respectful of, and intrigued by, what is unknown and unprecedented about his patients and who assumes that he or she might well be challenged by each patient to call upon something in him- or herself that is new and unprecedented in his or her experience” (Hoffman, 2009, p.1050).

Actually, in one study Shedler found that 99.4% of clinicians agreed that the SWAP allowed them to describe their patients in a meaningful way. Apparently, they did not feel forced into a Procrustean bed in which the uniqueness of their patient was lost or “desiccated” (Shedler, personal communication, recalls an instance in which a patient dressed up in animal suits and met with other “furries”. This patient’s therapist did not feel the SWAP items captured this important aspect of the patient). So, is it reasonable to think that the 1% of clinicians who did not feel that the SWAP allowed them to describe their patients are necessarily the more talented therapists the ones on whom we should put our "money"? Does finding the SWAP a meaningful instrument mean that one is less "profoundly respectful" of the patient or prevent the therapist from empathic immersion and appreciation of the subtle complexities of the patient's dynamics?

Hoffman cites Kazdin's (1998) point that one value of case studies is that they point to rare phenomena that cannot be evaluated in a research context. Hoffman extends this point by claiming that every case is not only "rare" but "unique" and argues that the logical implication of Kazdin's position is that any "evaluation in group research" is "impossible". If this were the case, a corollary implication is that the therapist cannot generalize from one patient to any other patient.

In short, in our view it is incorrect to suggest that psychoanalytic researchers are engaged in the "desiccation" of human experience because they are trying to measure aspects of it. Conducted in a clinically meaningful manner, such efforts and the research in which they are embedded are the best protection against authoritarian thought control precisely because this approach involves replicable empirical evidence rather than persuasive, charismatic appeals designed to strike a resonant emotional chord in others. These researchers do not rush to embrace a theory simply because it "speaks to" them (the advice proffered by Greenberg & Mitchell, 1983), but only if it passes muster based on the rigors of controlled testing. In this sense, it might be said that researchers operate closer to the reality principle than to the pleasure principle. Because they wish to reduce ambiguity by using controlled methods of observation does not mean they cannot tolerate it, as Hoffman would have us believe. The counter-claim would be to assert that clinicians cannot tolerate clarity and have a strong need to wallow in uncertainty.

Although he is careful to note that he is not claiming that research has *"no"* value (p. 1044), Hoffman appears to believe it has very little, that it is fundamentally a de-humanizing capitulation to power politics, and that it fails miserably in capturing the uniqueness of the individual. He (2009, p. 1047) approvingly quotes Cushman and Gilford (2000) who note that evidence-based approaches entail "abhorrence of ambiguity, complexity, uncertainty, perplexity, mystery, imperfection, and individual variations in treatment" (p.993)".

In short, Hoffman believes that systematic, quantitative research methods make a scant contribution to psychoanalysis. His dismissal of psychoanalytic research goes beyond Freud's reply to Rosenzweig's experiments. Freud said they were not necessary, but could do no harm whereas Hoffman believes they can be damaging.

Hoffman offers some support for his position by showing that certain questions are extremely difficult, if not impossible, to answer through research. For example, he cites the case of a 95 year old patient who is terrified of death and asks how can we distinguish between existential dread of death and neurotic death anxiety and tell the insurance company how many more sessions it will take to eliminate the neurotic anxiety. One need not claim that research can answer any and all clinical questions to accord it some value. As we shall show shortly, there are many other answerable questions that Hoffman fails to raise, questions that can never be answered by case studies.

A conformist versus a critical psychoanalysis

Hoffman seems to associate empirical research with a *conformist* rather than a *critical* psychoanalysis and asserts that the latter is essential for preserving "human freedom, for the dignity of the individual, for the meaningfulness of community, and for the sacrosanct integrity of every moment of experience" (p. 1064-1065). He hopes that "psychoanalysis can be newly empowered as a humanizing force in our culture and in the world" (p. 1065) and fears that denigration of the value of case studies will not allow this to happen. He claims that "The “consequential uniqueness” of each interaction and the indeterminacy associated with the free will of the participants make the individual case study especially suited for the advancement of “knowledge” ---that is, the progressive enrichment of sensibility ---in our field” (Hoffman, p. 1043). It is noteworthy that Hoffman puts the word “knowledge” in quotation marks and that he refers to the “progressive enrichment of *sensibility*” (italics added). Surely there is no argument with the proposition that each interaction is unique and consequential and that immersion in clinical cases can enrich our sensibilities. However, being content with enriched sensibilities does not bring us closer to knowledge concerning the relative importance of different therapeutic ingredients.

The charge of “authoritarian objectivism”

Hoffman’s quote from Orwell’s 1984 is but one example of his implication that there is an alignment between science and authoritarianism and thought control. The danger of privileging research over case reports, according to Hoffman, is that it is “flawed epistemologically” and “threatens to embody yet a new form of prescriptive, authoritarian objectivism” (p. 1045). Hoffman believes the main divide is between dangerous objectivism and the constructivism he favors. But, there need not be an inherent antagonism between these views. One can engage in dialectical constructivism as an approach to treating people. At the same time, one can also examine the results and consequences of negotiated realities, freer self-disclosures, a more democratic therapeutic atmosphere, and so on with respect to the patient’s functioning during and after therapy. In fact, one can compare it with that of patients who have been treated by more conservative therapists, who are given more to ritual than to spontaneity, and who rarely “throw away the book”. The very fact of seeking to determine which approach works better, on average, for which type of patient will actually serve to prevent the kind of authoritarianism of those who would hold up their own personal experience as a sufficient basis for what to believe.

These kinds of questions cannot be answered *only* by the analyst’s reflections. That is why many observers and critics of psychoanalysis have noted that the analytic situation is an arena par excellence for generating hypothesis but an inadequate setting for testing them. (We refer here to testing them in the form of the individual analyst drawing conclusions, as opposed to using analytic data (e.g., verbatim transcripts) to test various hypotheses, as Luborsky did with his symptom-context method). Spending so much time with patients and being told that one can do no little more than raise interesting hypotheses, impressions, and speculations, is understandably hard to accept. On the other hand, for those who favor a ‘constructivist’/‘perspectival’/ hermeneutic view in which there is a “consequential uniqueness” to each analytic dyad, one might think that it should be easier to accept the limitations of what an individual clinician can say in terms of generalizations concerning what is helpful in treatment. Yet, such clinicians do not appear content to leave it at “This is how I see it”; rather, they tend to present their findings as having some objective validity and as advancing knowledge, not just enriching “sensibilities”. Once such claims are made, expectations for evidential support beyond clinical judgments become legitimate as the therapist is not just describing his or her own subjective experience with patients but making knowledge claims..

It is not surprising that over the past decades analysts have been driven to make more modest claims about the validity of their theories and their treatment outcomes. The so-called ‘hermeneutic turn’ in psychoanalysis (some might call it a ‘retreat’) seems, in part, to be a result of the recognition that it is extremely difficult to establish the validity of interpretations if we stay entirely within the clinical situation. Instead of trying to show how we know that our interpretations “tally with what is real” in the patient we aim for a coherent, co-constructed, mutually negotiated narrative of the dynamics within the analytic dyad. Perhaps, as Frank (1961) suggested long ago, any reasonably coherent, plausible narrative will do. If so, we are forced to be modest about any theoretical claim based on interpretive work as possessing a greater degree of "accuracy" or validity than any other. However, there can be extra-clinical tests of competing hypotheses (e.g., Silverman, xxxx) which, because the studies control for confounding variables, offer a more secure basis for favoring some hypotheses rather than others. We should not have to elaborate the point that measuring variables and controlling for them helps to reduce the number of plausible, alternative explanations.

For what kinds of questions do answers from research deserve to be "privileged" over those offered by case studies?

Obviously, not *all* questions about treatment can be answered through research (e.g., Hoffman's 95 year old patient). However, there are many important questions that we can answer better through systematic research than through clinical cases studies, or, at least, we can see the extent to what we think we know from clinical work squares with what we can learn from research. We also need to recognize that some relevant questions can never be answered adequately if we rely exclusively on the case study method.

Here are a few examples where research has or can make a valuable contribution:

1. Is the optimal number of sessions per week different for patients with different diagnoses?
2. Do transference-focused, compared with non-transference, interpretations made to borderline patients result in faster and more stable improvement in relationships and decreased self-destructive behavior? And, how do the effects of transference-focused psychotherapy compared with those of rival treatments (e.g., dialectical behavior therapy)?
3. What is the relationship between the quality of the alliance at different points in treatment and outcome?
4. Do therapists adhering to different theories have different rates of success?
5. Which kinds of personality changes are more enduring when treated by psychoanalysis compared with other forms of treatment?
6. What is the relationship of therapist warmth and empathy to outcome?
7. Do therapists who adhere to a treatment manual generally achieve better outcomes than those who do not?
8. Under what circumstances does countertransference disclosure reverse a previously stalemated treatment? The literature is replete with case vignettes that purport to demonstrate that this often is the case. However, we have no base rate data, e.g., what percentage of the time does countertransference disclosure make any difference in the progress of treatment?
9. Do certain symptoms (e.g., stomach pains) get reported in particular thematic contexts rather than others? If so, this would provide some insight about the kinds of conflicts associated with particular symptoms. This is precisely what Luborsky (1996) did in devising the "symptom-context" method in which he compared the material just preceding and just following the report of a stomach symptom, compared with a control condition. This is the kind of study that could not be done using informal recollections of what patients said.
10. Luborsky's work on 'momentary forgetting' (Luborsky, 1977) and on the Core Conflictual Relationship Theme (CCRT) (Luborsky & Crits-Christoph, 1990), Bucci's research on 'referential activity' (Bucci & Maskit, 2007), and the studies by Safran & Muran (2000) on ruptures and repairs of the therapeutic alliance are examples of systematic empirical research using psychoanalytic data that have yielded valuable information of a kind not possible to extract from clinical case studies.
11. As a final example, consider the role accorded to transference interpretations.

Transference interpretations have long been assumed to be an essential element in psychoanalysis and in psychoanalytically-oriented psychotherapy. How would we ever know if this assumption is valid or the conditions (e.g., types of patients (e.g., level of object relations), the quality of the therapeutic alliance, etc.) under which it matters whether or not transference interpretations are part of the treatment? It is hard to imagine that we would ever know the answers to these questions merely on the basis of accumulated clinical experience.

Fortunately, the beginnings of an answer are to be found in a series of studies by Hoglend *et al*. (2006, 2007, 2008). These investigators studied 100 patients randomly assigned to a one year dynamic psychotherapy in which one group received no transference interpretations and the other group received a “moderate” level of transference interpretations. Each patient’s quality of object relations was assessed in a two hour audio taped pre-treatment interview that at least three clinicians listened to and rated. Several quantitative indices of change also were assessed. There was no main effect of conditions. That is, overall, it did not matter (either at termination or at 1 and 3 year follow-ups) whether the patients did or did not receive transference interpretations. However, there were significant interactions between group and quality of object relations. For example, on measures of level of “Psychodynamic Functioning”, for patients with high scores on “quality of object relations” there was no difference in the degree of “clinically significant change” as a function of whether the patient did or did not receive transference interpretations; 60% of those who received transference interpretations improved, compared with 55% of those who were not offered transference interpretations. In the case of patients with low scores on “quality of object relations”, less than 20% of the group that did not receive transference interpretations improved, compared with 40% in the group that did receive transference interpretations.

As with any study, there are limitations that point the way to future studies. Most psychoanalytic researchers, including Hoglend, are well aware that the results of their studies might not be generalizable to psychoanalysis or to long-term analytic therapy. In Hoglend’s studies, the treatments were manualized and although the therapists were trained to perform both treatments equally well, they might have been less comfortable when not permitted to use transference interpretations. Perhaps a “moderate” level of transference interpretations is not the optimal level to use with patients who have good object relations. Perhaps the results would have been different if the quality of the therapeutic alliance was assessed. Numerous other questions can be raised. The point is that this was, as the authors (Hoglend *et al*, 2006, p. 1740) say, “…the first experimental investigation designed to measure the effects of a moderate level of transference interpretations in brief dynamic psychotherapy”. As such, it deliberately sacrificed external validity for a higher degree of internal validity. It is an important start to answering key questions that have been around for more than a century. We do not want to overstate or understate the effectiveness of our clinical work. Imagine if there was no research challenging the claims of drug companies.

It should be apparent that the kinds of questions posed above are beyond the attention span or observing capabilities of an individual observer nor would the interchange of pooled information from many analysts would be sufficient to answer the questions. If we rely exclusively on the case study method, questions such as the optimal role of transference interpretation will be debated in the literature 100 years from now.

Can research offer any help to the clinician in the consulting room?

Hoffman is correct in noting that empirical research fails to offer the analyst any immediate help when he is working in the consulting room. However, the desirability of conducting psychoanalytic research on the process and outcome of treatment and on basic processes posited by psychoanalytic theory should not rest on whether or not it offers immediate help to the analyst. For example, if in the midst of a session, the analyst is conflicted about offering a countertransference based self-disclosure, there will be no body of research to serve as a *specific* guide. However, one could imagine that some day there might be *general* guidance from research findings regarding the circumstances under which different kinds of deliberate self-disclosures, with different kinds of patients, at different stages in treatment, facilitate or impede therapeutic progress. Similarly, if the analyst wishes to encourage the patient to overcome a reluctance to attend four sessions a week rather than three, the analyst potentially could be guided by the research findings on the relation between frequency of sessions and outcome. Of course, when it comes to processing the patient's associations in preparation for offering dream interpretations, research will be of no avail.

Let us grant that knowledge of the research literature would have absolutely nothing to offer Hoffman when he is in the consulting room trying to help a patient. But, what about the very fact that Hoffman is in his consulting room in the first place. How does he know that he is offering a service that helps even some of his patients? By his own judgment and that of the patient? How free are such judgments from cognitive dissonance? To what extent would they be seen the same way by independent observers? What about follow-up information? In other words, how stable and enduring are allegedly positive therapeutic outcomes? The point is that even if research were totally useless when it comes to guiding clinical work, research on outcome could tell us if it is even ethically proper to hold ourselves out to the public as effective healers.

The evidential value of case studies

From its inception throughout most of the 20th century, the psychoanalytic case study has enjoyed privileged status, vis a vis systematic empirical research, as the means of establishing and advancing psychoanalytic knowledge. Freud, who, as we know, started his medical career as a scientist, periodically voiced qualms about the evidential status of clinical case reports, but felt he had no alternative to relying on clinical observation and judgments to verify his theories. However, Freud clearly was aware of the inevitability of the influence of 'suggestion', and of theory-guided inferences and their compromised evidential value. For example, Freud (1916, p. 445) stated that "In so far as his transference bears a "plus" sign, it clothes the doctor with authority and this is transformed into belief in his communications and explanations", and "what is advantageous to our therapy is damaging to our researches" (p. 452).

At the start of his discussion of the case of Little Hans, Freud (1909, p. 104) wrote:

"It is true that during the analysis Hans had to be told many things that he could not say himself, that he had to be presented with thoughts which he had so far shown no signs of possessing, and that his attention had to be turned In the direction from which his father was expecting something to come. This detracts from the evidential value of the analysis; but the procedure is the same in every case. For psychoanalysis is not an impartial scientific Investigation, but a therapeutic measure. Its essence is not to prove anything, but merely to alter something. In a psycho-analysis the physician always gives his patient (sometimes to a greater and sometimes to a lesser extent) the conscious anticipatory ideas by the help of which he is put in a position to recognize and to grasp the unconscious material. For there are some patients who need more of such assistance and some who need less; but there are none who get through without some of it."

Freud also commented on the evidential status of case reports in his case of Fraulein Elisabeth von R. (Breuer & Freud, 1895). He wrote:

"I have not always been a psychotherapist. Like other neuropathologists, I was trained to employ local diagnoses and electro-prognosis, and it still strikes me myself as strange that the case histories I write should read like short stories and that, as one might say, they lack the serious stamp of science. I must console myself with the reflection that the nature of the subject is evidently responsible for this, rather than any preference of my own." (p. 160).

That Freud needed to remind the reader that he was a scientist and to let the reader know that he had to "console" himself suggests that he was troubled, though not deterred, by the story-like quality of case reports. Undoubtedly, one reason for this is that he wanted his assertions to bear the "serious stamp of science", even though they lacked its usual characteristics.

We know that most case reports consist of vignettes selected to support a hypothesis rather than being a complete and faithful account of what transpired. Thus, years after Freud's expressed his concerns, Anna Freud (1971, p. ix) implied a similar uneasiness when she noted that "We cannot help being conscious …of a conspicuous…dearth of…complete and adequately documented case histories." As Michels (2000) pointed out, a survey of the psychoanalytic literature from 1969—1982 that focused on the articles cited most frequently apparently failed to find any extensive case study reports (Klumpner & Frank, 1991). Other analysts, however, seem to feel that relying on selected case vignettes is fine, indeed preferable to full-length reports because they provide a more vivid account of the analytic work (e.g., Stein, 1988).

In relation to the issue of selectively regarding which aspects of which cases are found in the literature, Michels (2000) invites us to pay attention to the analyst's purposes in writing up a case and publishing it. He indicates that when the intention is to offer evidence for an analytic hypothesis about the meanings of some aspect of the patient's behavior, many observers believe it would be useful to have a tape and a transcript. On the other hand, as Michels (2000) notes, Galatzer-Levy (1991, p. 736), in a panel report of the Committee of Scientific Activity (of the American Psychoanalytic Association), comments that the preference for verbatim data is "scientism", "…the irrational veneration of what appears scientific rather than using scientific methods as tools." He states "Abandoning narratives would deprive us of the richly informative narrator's perspective". This view represents an unnecessary choice. It need not be 'either-or'. Obviously, the narrator's perspective can be "richly informative", and would be even more informative if it was accompanied by a record of the thoughts and feelings the analyst experienced during the sessions on which the narrative is based. At the same time, the "richness" would be enhanced by *also* having the verbatim material for others to study in a systematic fashion. In fact, comparing the analyst's narrative with what might emerge from a detailed study of the original data by independent observers could be quite illuminating and more "richly informative" than either source of data alone. Such an approach would reduce the common limitations of case studies: 1. Deliberate distortion of case material and/or facts in the patient's history in the service of presenting a more compelling set of assertions; 2. Unwitting distortion or selective memory of facts and/or clinical data in the service of offering a more persuasive case or as a result of countertransference reactions; 3. Deliberate disguise of the patient's identity that results in the alteration of clinical data and biographical facts about the patient's history. This makes is difficult for the reader to judge the evidential value of the data.

The problem of "confirmatory bias"

A systematic, empirical approach might shed light on the issue of biased weighting of clinical evidence. In this regard, one of us participated in a research project on clinical evidence in which several analysts studied the verbatim transcripts of numerous analytic sessions. The group, organized and led by Benjamin Rubinstein, met regularly. We started by reading the transcripts of the first five sessions. Any time a member of the group had a hypothesis to offer, we stopped and recorded the hypothesis and the observations on which it was based. In subsequent meetings, we read transcripts of randomly selected subsequent sessions. When a group member felt there was evidence, either in favor of, or against, a given hypothesis, we stopped and rated the strength of the evidence. Two noteworthy findings emerged from this procedure. First, 98% of the ratings were in the positive direction, meaning that we rarely regarded a hypothesis to have been disconfirmed by the clinical material. Second, when we compared the 'strength of evidence' ratings of the person who had us stop to rate the evidence for a given hypothesis and compared that rating to the average the rating of the other group members (i.e., those who did not call attention to clinical material), the group rating was lower.

What this analysis suggests is that the analyst who felt there was evidence for a hypothesis (which did not necessarily have to be the one he himself proposed originally) thought the evidence was stronger than did his colleagues. In short, there was an indication of what we might call a "confirmatory bias", defined above, and expressed in our group by the tendency to give more weight to evidence than other colleagues feel is warranted. Another noteworthy finding is that it was quite rare (i.e., less than 5% of the time) for anyone to find negative evidence of any initial hypothesis. This finding is somewhat ambiguous in that it could reflect a confirmatory bias or the extraordinary clinical acumen of the clinicians! Extrapolating from these findings to the clinical situation, it is likely that 1) we rarely regard our initial hypotheses as disconfirmed or as not supported by further clinical observations, and, 2) we give greater weight to apparently "confirmatory" evidence than is warranted. It seems reasonable to regard this as a limitation of the case study method. At the very least, this kind of 'confirmatory bias' suggests room for improvement in the processing and reporting of case material.

Suggestions for improving the quality and bolstering the evidential value of case studies

Even if one wants to maintain that case studies are all that we need in psychoanalysis, Hoffman's paper is silent on the question of the adequacy of the typical case study (or clinical vignette) found in the analytic literature. If he wants to persuade us that such studies are to be preferred over research studies as the basis for psychoanalytic knowledge, one would have hoped that he might spell out the criteria for case studies that should be met. We can agree that case studies can offer ideas for new ways of thinking about the work and about patients, e.g., might serve, as Kohut's writings did, to legitimize the use of and emphasis on empathy as an important curative factor. They can allow one to re-think past and ongoing cases with the benefits of insights gleaned from reading case studies. Case studies can be of value in these ways even if they are distorted or falsified in certain respects, as in Kohut's (1979) two cases of Mr. Z.

Some authors have argued that properly conducted case studies can contribute to knowledge, but, unlike Hoffman, they spell out some criteria for acceptable case studies. Wile (2007), for example tells us that cases with a "rich case record" can be very valuable. Among the criteria he proposes are "Recordings of treatment sessions: verbatim transcriptions of audio or video recordings are a particularly strong source for grounding your inferences." He also advocates "Session-by-session assessments: repeated measurements of the client's problems, goals, symptoms, together with evaluations of sessions and the strength of the client-practitioner relationship"; "Outcome assessments", including both qualitative and quantitative factors; "Post-treatment interviews". In addition, he urges a careful review of all the available information. In other words, He is encouraging a more disciplined approach to clinical case studies than is typical. It is noteworthy that nowhere does Hoffman offer any suggestions for enhancing the degree of confidence one could place in inferences based on case studies. This is an especially thorny issue given what appears to be the fairly widespread belief that many clinical vignettes in the literature are composite pictures and some are altered in significant ways or even largely fabricated, partly to preserve the patient's privacy and, it seems, partly to make a theoretical point by selectively focusing on a segment of material.

Other commentators also have offered methods for improving the evidential value of case studies. As in Wile's (2007) paper, these methods rest on a distinction between clinical case studies as ordinarily presented in the literature (i.e., selected vignettes) and a more systematic study of audio-taped and videotaped sessions. For example, recorded sessions can be rated by independent observers on a wide variety of measures, e.g., changes in the patient's affect or self-reflective capacity over time, changes in the quality of the therapeutic alliance, changes in the intensity and frequency of symptoms, etc.

Case studies should be accorded more evidential value to the extent that they demonstrate the following characteristics:

1. The ratio of theory to data is reasonable, i.e., there is not an excessive amount of theory superimposed on some fragment of data.
2. Observation is clearly separated from inference in the case report/.
3. Alternative and rival hypotheses are considered seriously and the reasons for rejecting them presented.
4. The report is relatively free of jargon.
5 The report illuminates a phenomenon, justifies a particular technical approach or innovation, or argues cogently for an improved conceptualization of a familiar phenomenon.
6. Verbatim accounts are included.
7. The case formulation is internally consistent and coherent.
8. The case report is sparse and tentative with respect to etiological claims.
9. Issues of generalizability are considered carefully.
10. The report should not be based on a fictionalized case or a composite based on several cases.
11. It should be demonstrated that the inferences grew out of the material and were not imposed prematurely on the clinical observations, even if the vignette is selected to advance a particular point of view.
12.. Caution is evident in cause-and-effect types of claims regarding the patient's dynamics.
13. Caution is evident in cause-and-effect types of claims regarding childhood causes of current problems.
14. The author's report reflects an awareness of having read and absorbed the cogent points in Paul Meehl's classic paper, "Why I Do Not Attend Case Conferences" (the details of which I will not provide here).
15. There is independent confirmation of some of the claims made.
16. There is follow-up information on the case that bears on some of the assertions put forward.
17. The author recognizes the issue of base rates. For example, if it is alleged that a stalemate in a lengthy analysis was broken by a countertransference self-disclosure, the author should inform us of how often such self-disclosures did not seem to make any difference and perhaps offer some hypotheses in this regard. We also would need to know how often a stalemate is broken in the absence of countertransference based self-disclosure.

Very few, if any, case studies meet all, or even most, of these criteria. If they did, one could make a stronger case for their evidential value. As Kazdin (2001) notes, a serious commitment to patient care should include a recognition of the limitations of informal clinical judgment and the desirability of using supplementary methods of evaluation (e.g., Clement, 2001). Wakefield (2007), for example, provides a convincing example of Freud's distortion in the case of Little Hans when Freud incorrectly claimed that Little Hans confirmed that the giraffe represented his father. And, this is distortion not based on memory but on a misreading of the case record. This famous case highlights the distinction between the accuracy or probative value of case material and the influence it can have on generations of clinicians.

Edelson (1984) mounted a spirited defense against Grunbaum's (1977) challenge regarding the evidential merit of the case study method. Grunbaum (1977), it will be recalled, pointed to the fallibility of memory, the selection bias of the analyst, and the factor of suggestion as rendering the data obtained in the analytic situation irrevocably contaminated and unusable as probative evidence for analytic claims. In the face of these difficulties, Edelson (1984) believes that many analysts have a sense of futility in meeting the standards for evidence. It is Edelson's impression(1984, p. 157) that "This sense of inadequacy, and the despair that goes with it, may in some cases at least translate into abandonment of scientific for hermeneutic goals."

Edelson (1984, p. 158) presents more than a dozen specific suggestions for strengthening the probative value of case studies. His suggestions include: "1. Seek falsification rather than confirmation in case studies"; "3. Use causal modeling and statistical controls …"; "5.a) "Minimize suggestion through a disciplined use of psychoanalytic technique", b) Predict responses by the analysand to an interpretation that have not previously been manifested and that are not suggested in the interpretation". Few, if any, clinicians have done the needed work urged 25 years ago by Edelson (1984, p.160), whose last sentence in his book was that doing such work "is just one of the responsibilities that goes with being a psychoanalyst."

Intensive study of singles cases

Messer (2007, p. 55) describes the Pragmatic Case Study (PCS) method as a way to "…enhance the quality and rigor of knowledge gained from psychoanalytic single case studies." In line with the points made above, Messer (2007) notes the limitations of the typical case presentations of clinical vignettes. Such vignettes (1) rely exclusively on the therapist's notes or memory, (2) they are selected by the therapist alone, (3) they are interpreted in "…terms of reigning theoretical orthodoxy", (4) there is insufficient context to enable the reader to refute or accept the therapist's view of the case, and (5) the therapist is the only one that has access to the data. One might add that we only have the analyst's opinion of the patient's degree of improvement.

The PCS method outlined by Messer (2007) requires that the session (s) be videotaped or audio-taped, or that at least be detailed notes made during or immediately after each session. A third party could sample the tapes to see if the therapist was being selective. Alternative hypotheses would be seriously entertained rather than using the case report only to illustrate one's preferred theory. Providing more information and more context would also allow for alternative readings of the material. With regard to the problem of restricting the data to the therapist's report, Messer (2007) recommends that we include subjective and objective measures. For example, the patient could periodically complete standardized self-report measures (e.g., anxiety and depression scales).

Clement (2007) presents an extremely detailed single case study which includes estimates of the degree of change in the patient, compared with the therapist's other patients with the same problem. He selected a patient who presented with OCD, panic disorder, fear of flying, and nightmare disorder. She had a history of separation anxiety disorder. This 30 year old patient was in treatment for 103 sessions over a 2 1/2 year period. She was evaluated a year later and was found to have maintained her gains. Clement (2007) was able to calculate a treatment effect size of 3.95 (which is rather substantial) and a 3.81 effect size one year after termination (also quite impressive). He was able to compare this outcome with effect sizes of 1.12 to 1.56 in meta-analyses of OCD treatments that used randomized clinical trials. He also was able to compare this outcome with the average effect size (1.65) for the other OCD patients (N = 64) in his 40 years of practice. Thus, by these numbers the patient's improvement was quite significant.

Speaking in clinical rather than 'effect size' terms, Clements (2007) reports that of the 64 cases, he eliminated the 23.4% who did not continue for more than three sessions and one that was still in treatment. He is able to tell us that of his sample, 2% were "much worse" at termination, 35% showed "no change", 38% were "improved" and 25% (including the present patient) were "much improved", yielding an overall improvement rate of 63%.

Clement's (2007) calculation of effect sizes borrows from the method introduced by Smith and Glass (1977). He asked this patient to complete a self-report Adult Problems Checklist at intake and again in the 14^{th}, 69^{th}, and 103^{rd} session. The patient served as her own control so the effect calculated effect size is based entirely on her self-report. Although the systematic way that

Clement gathered information can be considered a step forward, the absence of independent observers' evaluations of outcome remains a limitation, especially because we do not know the effects of giving the patient a "therapy report card" after each self-evaluation. In this regard, Clements (2007) reports a .52 correlation, for his 683 patients over 26 years of practice, between degree of improvement at termination and number of sessions. Such a finding might be partly due to a cognitive/emotional dissonance effect.

In any case, Clement's (2007) approach allows him to tell prospective patients their chances of improving in treatment with him, relying on his data base of previous patients' self-reports rather than on the literature on efficacy studies from which one can hardly generalize (recall that Hoffman objected that research does not control for who the therapist is). From a pragmatic and ethical perspective, this is information patients are entitled to but that few therapists are able to provide.

It would seem that the kind of record-keeping that Clements does combined with the kinds of suggestions that Messer (2007) and we have presented would allow for a collection of intensively studied cases that would yield improved information concerning therapeutic outcome. It will come as no surprise to note that Clements' approach is a cognitive-behavioral one. We cannot realistically expect that analysts, or, for that matter, most therapists of any persuasion, would be inclined to gather and organize such detailed information on their cases.

Alleged protection against bias in clinical case studies

Hoffman tries to counter the charge of biased reporting by analysts by stressing the ambiguity of clinical data as a protection against bias. He claims that the fact of ambiguity makes clinical data "relatively unmanipulable" and that the ambiguity "ensures the openness of the "data" to critical review and to multiple interpretations" (p. xcxxx). The fact that other analysts can and do suggest alternative formulations and/or technical interventions does not ameliorate the problem of initially biased reporting or the constraints on the kind of knowledge claims that legitimately be made on the basis of case studies. Research data also can suggest alternative interpretations. One difference is that in the case of research data one can then go on to test the relative merits of the different interpretations. The clinician might counter that subsequent work with the patient or new patients offers the opportunity to select among alternative interpretations and compare the clinical evidence favoring one or another inference. The problem, however, is that there is no independent check on the accuracy of the inference or even whether other clinicians would see it the same way. All this is okay if the main goal or virtue of the clinical case study is to stimulate the clinician's thinking and is not intended to make claims about lawful regularities in mental life.

Hoffman correctly notes that changes in attitudes and behavior on the part of analysts have not been based on systematic empirical research. For example, the acknowledgment of the "intersubjective nature of psychoanalytic data" and the increasing "democratization of the analytic relationship" are clearly not the result of research We agree and would add that this can be said of psychoanalytic thinking for more than a century. The question is are changes brought about by shifts in cultural attitudes, philosophical or political values, charismatic theorists, and so on, to be considered accretions to knowledge and thus "progress" (e.g., improved outcomes or more valid theories of mental functioning) or are they largely shifts in prevailing fashions?

A variety of "good" ways of being with patients

Hoffman asserts that there are "multiple good ways to be, in the moment…", meaning in trying to help one's patients.(p. 1043). This sounds like a reasonable claim. No one today claims that there is a singular, correct technique. However, this view leaves aside the researchable question of whether some ways of being "good" are better than others. Here we would have to specify criteria for "good" and gradations of "good" as well as specifying when the encounter is no longer "good" but has turned "bad". Should this just be a judgment call that Hoffman or others can make by reference only to qualitative data as garnered by clinical impressions or would it imperil the sanctity and ecological purity of the analytic situation if transcripts of sessions were rated for various factors (e.g., new memories, degree of affect expressed, quality of the alliance, degree of resistance, quality of self-reflection, etc,)? Would this really constitute a "desiccation" of human experience or could we say that it is an attempt to capture aspects of uniquely human experiences for the ultimate purpose of facilitating the fuller flowering of human potential,? As noted earlier, to suggest that to measure some aspects of human experience is, by that very act, to destroy the experience does not seem valid. It is hard to see how trying to learn something in a systematic way by studying audio-recorded sessions in any way detracts from the analytic experience of the therapist or patient. Such studies might tell us that some ways of being with a patient are more beneficial than others.

Hoffman supports his view by analogy to the Rorschach, pointing out that it is an instrument used to assess reality testing. In the Rorschach, there are many acceptable responses to the ambiguous inkblots, acceptable in that they do not violate the shared consensus about the realistic properties of the inkblots. Hoffman implies that one supposedly cannot say that any response that conforms to the properties of the inkblot is better than any other. But, his analogy does not hold. To stay within his analogy, there are some responses that include color, shading, movement, and/or texture as response determinants and some that don't, even if the same percept is given. Some responses are offered quickly, others are delayed. Individuals vary in how many original versus 'popular' responses they give as well as the total number of responses they give. Although these variations do not in and of themselves speak to the issue of reality testing, they are interpreted by clinicians as "better" or "worse" in terms of what they are thought to reveal about personality functioning (e.g.., the modulation of affect, defensive style, regression in the service of the ego, etc.).

Going back to the clinical situation, one can say that there are multiple ways of fostering a therapeutic alliance and promoting the patient's sense of safety and analytic trust. However, some of these ways might be better than others. There are many variations in the timing, depth, and dosage of interpretations, whether they focus mainly or only on the 'here-and-now' or whether they involve the historical past, and so on. Many such interpretations might be comparable with respect to the patient feeling understood or feeling that the interpretations are plausible and convincing. However, might we not learn something beyond the clinician's impressions, or different impressions by different colleagues, as to which line of or frequency of interpretations is more beneficial in moving the analysis forward. Of course, we would have to specify the criteria for what we mean by "beneficial". In this regard, consider the work on transference-focused psychotherapy (TFP) with borderline patients (Clarkin, Yeomans & Kernberg, 1999; Kernberg, xxxx). Such an approach seems to increase self-reflective capacity and is related with a decrease in self-injurious behavior. Comparisons to schema therapy (Arntz, xxxx) and dialectical behavior therapy (DBT, Linehan, 1993) are necessary to determine which, if any, of these three treatment approaches to borderline personality disorder are, on average, superior, and to what degree. Obviously, there will be disagreements among adherents to these approaches, each claiming that outcomes that support a rival approach are methodologically flawed. But, these kinds of criticisms can be responded to in the next study. However, aside from comparisons among approaches is the question that Kernberg and his colleagues have not yet tackled.

So, if we want to know whether, generally speaking, transference interpretations are useful for borderline patients and we do not wish to rely entirely on the recollections and informal impressions of therapists treating such patients, does that really mean that we cannot tolerate ambiguity, uncertainty, imperfection, and so on, or that we would not be sensitive to differences among patients in terms of their receptivity to transference interpretations?

Can avoidance of research or research findings harm patients and their families?

Hoffman argues that research does not help the clinician and that it is sufficient to rely on the clinical case study. At least one glaring counter example to this claim is the many years it took to de-pathologize homosexuality (Friedman, 1988; Friedman & Downey, 1998). Without the empirical research of Evelyn Hooker (1957, 1959) and others (ref) showing that homosexuals showed no more evidence of psychopathology than heterosexuals, it would have taken much longer for this realization if we did not value non-psychoanalytic data.. Much needless suffering could have been avoided if analysts did not remain wedded to their pet theories. The same can be said regarding the history of theories of autism that emphasized ‘refrigerator mothers’ and of theories of schizophrenia. that focused on ‘schizophrenogenic’ mothers. It seems fair to say that although both clinical/theorerical thinking and research-based scientific thinking can change in response to evidence, for a variety of reasons, the former generally is far more slowly self-correcting than the latter.

Research does not control for who the therapist is

Hoffman complains that research does not control for who the therapist is. We are not entirely clear what he means by this. We assume his point is not that there is a failure to take into account therapist variables for he knows that there are scores of studies that do include that element in the research design. If he means that none of the research findings could apply to any one therapist, including himself, because he himself was included in the studies, it would then suggest he is once again emphasizing the “consequential uniqueness” of each analytic dyad. One logical implication of this position is that what he learned about psychopathology and the analytic process with one patient has little, if any, generality to his work with his next patient. That would mean he has no store of cumulative knowledge on which he can draw. We doubt that he really means this, but this is a point that needs clarification.

In a related vein, we wonder if Hoffman is saying that the only research results he would have any faith in would be ones that were based on single case studies of therapists, e.g., tracking an individual therapist’s rate and degree of “success” and “failure” across a large sample of his patients. We suspect that even with generous government funding, including high fees for the service providers, few, if any, analysts would volunteer to participate in such a study. Yet, would we not agree, in principle, that part of accountability includes a moral component, one reflection of which is the willingness to have one’s results with patients be open to scrutiny by others? It strikes us as contradictory to be a staunch champion of human dignity, of human rights, and of the democratization of the analytic situation and yet not ever urge an investigation of how well the patients’ welfare is served by therapy. In any case, Clement’s (xxxx) approach is one that does take into account who the therapist is..

Constructivism versus objectivism

Hoffman has devoted many years to furthering his cause of a constructivist view of the psychoanalytic situation. We understand that his formulations strike an emotionally resonant chord in many analysts, perhaps in significant measure because they call for a democratization of the analytic relationship and give permission to analysts to “throw away the book”, to emerge from the shackles of analytic orthodoxy and to become more spontaneous and less ritualized in their interactions with patients. But, Hoffman is not simply offering his views simply as his subjective opinion but in the context of forceful arguments that it is a position superior to what he calls an "objectivist" view, though he does not define that term in his article. Is it not fair to say that Hoffman is being an "objectivist" in asserting that a "constructivist" view is not merely an alternative view but a better one, both conceptually and practically (i.e., that it is more likely to help patients)? How would we ever find out if it makes any real world difference whether one adopts one position or another unless we were willing to put the issue to an empirical test? This would mean a properly selected sample of self-identified "constructivists" and "objectivists" treating a large enough group of patients to provide the statistical power to detect clinically meaningful differences. The results would have to be replicated a few times and the studies would have to avoid "allegiance effects" by using theoretically "neutral" investigators. In the absence of such a summer camp-type "color war" competition of this kind, which is unlikely to ever take place, it strikes us as at least bordering on arrogance to imply that the "objectivists" are the 'bad guys', the ones more inclined toward authoritarianism, the ones who would be more apt to find the SWAP a clinically meaningful instrument for rating their patients and, therefore, presumably be less gifted therapists than those who felt the SWAP items did not offer a good description of their patients. Recall that it was the latter group that Hoffman said he would put his money on.

Hoffman equates empirical research with "objectivism", which he contrasts with the "constructivist" view he has been promoting for years. He seems to assume that there is an inherent contradiction between maintaining a constructivist perspective and encouraging empirical research, whereas we do not see any such incompatibility. To cite but two concrete examples: 1) one can examine the audio recordings of patients treated by analysts who identify themselves as "constructivists" and study patient-therapist interactions with regard to any number of possible relationships (e.g., the correlation between ratings of therapist "warmth" and outcome), 2) one can investigate the relationship between the ability of clinicians to engage in self-analytic work and changes in their attitudes toward the patient (e.g., Schlesinger & Wolitzky, 2002). Although Hoffman claims that he does not object to all research, the thrust of his arguments suggest that his version of "constructivism" is closely linked with a negative attitude toward measurement and quantification on the grounds that such activities cannot fully capture the uniqueness of the individual and that they compromise human dignity.

While insisting on the sacrosanct nature of the patient-therapist relationship is a morally laudable position, it often entails a visceral objection on the part of seasoned practitioners to virtually all attempts to study the nature of that relationship other than by anecdotal clinical reports. Not surprisingly, this is one reason that the bulk of the psychotherapy research literature consists of studies in which the therapists were novices, thus limiting the generalizations that can be made.

CONCLUDING COMMENT

Most observers probably would agree that empirical research and case studies each has certain advantages and certain liabilities and that they can complement one another. From this apolitical perspective, one can ask what each approach can contribute to our understanding.

The advantages of the case study method are that: 1) it enables us to study rare phenomena, 2) it generates insights and hypotheses about personality dynamics that are not readily elicited in other situations, 3) it suggests different kinds of interventions, and, 4) it can disconfirm certain hypotheses by finding instances that run counter to a theory. Perhaps the main limitation of the case study method is that it does not offer a good way of choosing among alternative hypotheses. In addition, the data are unreliable. They often are fictionalized, composite, or selectively remembered or used accounts designed to make a particular point. There is no opportunity for others to examine the data on which the clinician's conclusions are based. Collections of clinical anecdotes do not add up to a reliable body of data. As Fonagy (xxxx) quipped, among analysts, the plural of anecdote is "data"!

The advantages of systematic empirical research are that its controlled nature enables us to reduce speculative inferences since it is easier to rule out alternative explanations. In this sense, the data and inferences from them are less likely to reflect subjective bias. The main disadvantage is that the ecological validity of the phenomenon being studied might be excessively sacrificed in order to ensure the internal validity of the research design. To the extent that this is the case, external validity is limited. Perhaps the best compromise is to use clinical data that are collected in the naturalistic setting for a more systematic study than is possible for an individual therapist. As to Hoffman's objection that the individual therapist is left out to the equation, that need not be the case. For example, it is perfectly possible to study therapist differences in success rates across a sample of patients. We would then be able to answer such questions as: 1) do certain matches or mismatches in personality styles, values, attachment styles, etc. make a difference with respect to treatment outcome?, 2) do some therapists have consistently better or worse outcomes than others? To pursue the answers to questions such as these need not in any way "desiccate" human experience, as Hoffman would have us believe. Nor would this line of investigation be contrary to the acknowledgment of the "consequential uniqueness" of each analytic dyad.

Thus, arguing in terms of clinical case studies *versus* systematic empirical research is overly simplistic and fruitless. In order to decide which kind of approach and which kind of "evidence" to "privilege" we need to know the *nature of the question that is being asked*. The therapist's information processing capacity has limits, as does introspection, and freedom from bias. This obvious fact does not denigrate the therapist any more than saying that one can see more through a microscope than with the naked eye. Thus, even with the pooling of observations and memories across many therapists we could not answer questions of etiology or of whether, on average, more frequent sessions, longer treatments, degree and frequency of therapist self-disclosure, and a host of other issues relevant to psychoanalytic treatment without some objective measures made by outside observers. A physician cannot be expected to determine the patient's cholesterol level without a blood test.

The long reigning hegemony of the case study has made us vulnerable to some rather harsh reactions by scientifically-minded colleagues, despite the fact that we are beginning to offer good answers to our critics (e.g., Shedler, 2010). For instance, in his prefatory editorial to the blistering critique of clinical psychology recently published by Baker, McFall, and Shoham (2009), Mischel (2009, p. i) states that "The disconnect between much of clinical practice and the advances in psychological science is an unconscionable embarrassment for many reasons, and a case of professional cognitive dissonance with heavy costs". He also approvingly quotes Paul Meehl: "in one of his last public speeches, memorably noted that most clinical psychologists select their methods like kids make choices in a candy store: They look around, maybe sample a bit, and choose what they like, whatever feels good to them" (p. i). Although the Baker *et al.* and Mischel views are too harsh an indictment of the clinical enterprise, this need not lead to a wholesale dismissal of the importance of empirical research.

Lest Meehl's statement seem like totally unfair ridicule, consider the advice proffered by Greenberg and Mitchell (1983), in their now classic text. When it comes to embracing a theoretical point of view, these authors advised the practitioner to adopt whatever theory "speaks to you", that is generates the greatest emotional resonance. This position is not even balanced by a suggestion that one also read the relevant research literature to see to what extent the theory one resonates to has received at least some empirical support. In a similar vein, Mitchell (19xx) characterizes those who are concerned with the issue of evidence as suffering from what he sarcastically dubbed the “Grunbaum Syndrome”, allegedly a pathological state of mind, named after the philosopher of science who pointed out the epistemological liabilities of psychoanalytic theory (Grunbaum, 1984). It is one thing to say that in the immediacy of the clinical situation, it is probably inevitable that the analyst will process the patient’s material through the theoretical lenses that are most meaningful to the analyst. This probably is the only way one can proceed. But, to hold up such an attitude as an ideal and to disparage research as useless will not advance psychoanalysis.

References [incomplete list]

Baker, McFall, and Shoham (2009). xxxx.

Breuer J. & Freud S., *Studies on Hysteria* (1892 [1893]-1895), Chapter IV. *Standard Edition,* 2

Breuer J. & Freud S. (1892-95). Studies on Hysteria. *SE,* 2

Bucci W. & Maskit B. (2007). xxxx.

Clarkin J.F., Yeomans F. & Kernberg O.F. (1999). *Psychotherapy for Borderline Personality*. New York: Wiley.

Clement (2001). xxxx.

Clement (2007). xxxx.

Cushman & Gilford (2000). xxxx.

Edelson M. (1984). *Hypothesis and Evidence in Psychoanalysis*. Chicago: Univ. of Chicago Press (trad. it.: *Ipotesi e prova in psicoanalisi*. Roma: Astrolabio, 1986).

Hooker E. (1957, 1959) and others (ref xxxx.

Frank J.D. (1961). *Persuasion and Healing: A Comparative Study of Psychotherapy*. New York: Schocken Books (rev. ed.: Baltimore: Johns Hopkins University Press, 1973).

Freud A. (1971). xxxx.

Freud S. (1916). p. 445 xxxx.

Freud S. (1908 [1909]). Little Hans. xxxx.

Friedman (1988). xxxx.

Friedman & Downey (1998). xxxx.

Galatzer-Levy (1991), p. 736 xxxx.

Gill M.M. (1994). *Psychoanalysis in Transition: A Personal View*. Hillsdale, NJ: Analytic Press.

Greenberg J.R. & Mitchell S.A. (1983). *Object Relations in Psychoanalytic Theory.* Cambridge, MA: Harvard Univ. Press.

Grunbaum (1977). xxxx.

Grünbaum A. (1984). *The Foundations of Psyhoanalysis. A Philosophical Critique*. Berkeley, CA: Univ. of California Press.

Hoffman I.Z. (2009). Doublethinking our way to “scientific” legitimacy: The desiccation of human experience." *Journal of the American Psychoanalytic Association*, 57, 5: 1043-1069.

Hoglend *et al*. (2006, 2007, 2008). xxxx.

Kazdin (2001). xxxx.

Kazdin (1998). xxxx.

Klumpner & Frank (1991). xxxx.

Kohut H. (1979). The two analyses of Mr. Z. *Int. J. Psychoanal.*, 60: 3-27.

Linehan M.M. (1993). *Cognitive-Behavioral Treatment of Borderline Personality Disorder*; and *Skills Training Manual for Treating Borderline Personality Disorder* New York: Guilford.

Luborsky L. (1977) momentary forgetting’ xxxx.

Luborsky L. (1996). *The Symptom-Context Method: Symptoms as Opportunities in Psychotherapy*. Washington, D.C.: American Psychological Association.

Luborsky L. & Crits-Christoph P. (1990). *Understanding Transference: The CCRT Method.* New York: Basic Books.

Luborsky L., Diguer L., Seligman D.A., Rosenthal R., Krause E.D., Johnson S. *et al*. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology-Science and Practice*, 6, 1: 95-106.
Luborsky L., Rosenthal R., Diguer L., Andrusyna T.P., Berman J.S., Levitt J.T., Seligman D.A. & Krause E.D. (2002). The Dodo bird verdict is alive and well – mostly. *Clinical Psychology: Science and Practice*, 9, 1: 2-12 (Commentaries [pp. 13-34]: D.L. Chambless; B.J. Rounsaville & K.M. Carrol; S. Messer & J. Wampold; K.J. Schneider; D.F. Klein; L.E. Beutler).
Messer S.B. (2007). Pragmatic Case Study (PCS , p. 55 xxxx.
Michels R. (2000). xxxx.
Mischel (2009). xxxx.
Mitchell S. (19xx). xxxx.
Orwell G. (1949). *1984 (Nineteen Eighty-Four)*. London: Secker and Warburg.
PDM Task Force (2006). *Psychodynamic Diagnostic Manual (PDM).* Silver Spring, MD: Alliance of Psychoanalytic Organizations [see Internet: http://www.pdm1.org].
Safran J.D. & Muran J.C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. New York: Guilford Press.
Schlesinger & Wolitzky (2002). xxxx.
Shedler J. & Westen D. (2007). The Shedler-Westen Assessment Procedure (SWAP): Making personality diagnosis clinically meaningful. *Journal of Personality Assessment*, 89: 41-55.
Shedler J. (2010). The efficacy of psychodynamic psychotherapy. *American Psychologist*, 65, 2: 98-109. Internet edition: http://www.apsa.org/portals/1/docs/news/JonathanShedlerStudy20100202.pdf.
Silverman, xxxx
Smith and Glass (1977). xxxx.
Stein, 1988. xxxx.
Wakefield J.C. (2007). xxxx.
Westen D., Morrison Novotny K. & Thompson-Brenner H. (2004). The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 130: 631-663. Also in: PDM Task Force, 2006.
Wile (2007). xxxx.