We share Dr Sackeim’s concerns about the effects of electroconvulsive therapy (ECT) on autobiographical memory, and at no point do we deny that ECT can cause retrograde amnesia. From what our patients tell us, it certainly occurs. However, we are also aware that both the passage of time and depression affect autobiographical memory function. Very simply, the issue is that we do not actually know what are the nature, incidence, and extent of retrograde amnesia after ECT nor how long it persists. We do ourselves and our patients a disservice if we believe and extent of retrograde amnesia after ECT nor how long it persists. We do ourselves and our patients a disservice if we believe that under discussion: the Columbia University Autobiographical Memory Interview (CUAMI) and its short form (CUAMI-SF). If we do not know what exactly the test is measuring, how well it measures what it is intended to measure, how consistent the test is and free from measurement error, and what the normative results are for populations of interest (e.g., healthy controls, depressed non-ECT, depressed + ECT) with different demographic features (e.g., age and education), then test results cannot be sensibly interpreted. In fact, tests with diagnostic aspirations (i.e., detecting retrograde amnesia) should not really be used until such data are readily and transparently available.

Dr Sackeim has provided a nice account of pioneering work on the effects of electrode placement, dosage, and pulse width on ECT outcomes. In relation to these studies, he has described work on developing the CUAMI and the CUAMI-SF along with various studies in the ensuing 2 decades in which he has applied these instruments to try to assess retrograde amnesia. These studies show that CUAMI scores of percent recall consistency (compared to baseline performance, irrespective of how poor that might be) often differ between treatment groups with, for example, right unilateral and ultrabrief pulse ECT being associated with better percent recall scores. Additionally, he presents a series of interesting, but post hoc, correlation studies of CUAMI results with a range of other parameters (e.g., modified Mini Mental State, electroencephalographic changes, subjective memory reports) that were all performed after the “development” of the instrument.

Oddly enough, “reliability” is being confused with “validity”. To take just one example, as a proof that the CUAMI is “highly reliable”, Sackeim lists (p. 180) a post hoc correlation with a one-item self-report of cognitive function. Even if such a correlation was hypothesized before the original study was conducted, it would not support the CUAMI’s reliability, that is, the degree to which the instrument is producing similar results under similar conditions. The reliability of an instrument cannot be established by its correlation with another instrument or, in this case, with a single question item. Interestingly, even the authors themselves doubt the reliability of this single question item: “The GSE-My consisted of only a single item, perhaps limiting reliability of assessment.” Thus, according to Dr Sackeim’s rationale, an instrument’s ability to produce identical results under similar conditions can be demonstrated by its association with a different instrument with limited reliability! Dr Sackeim’s paper abounds with several other equally unusual arguments about the assessment of reliability and validity.

Let us be very clear here. Despite all this fine work and Dr Sackeim’s somewhat Procrustean arguments, at no point have there been any formal peer-reviewed studies to assess the actual validity and reliability of the CUAMI or the CUAMI-SF instruments before they were applied in clinical trials and other studies. As for normative data, all we have to date are one-off baseline CUAMI scores for 16 nondepressed controls reported in 1995 and baseline CUAMI-SF scores for 24 nondepressed controls reported in 2007. Recall consistency, the main “result” of these instruments, was strangely never assessed in the former study and not published for the latter group, although it was reported to be substantially better for the healthy controls than for patients treated with bilateral ECT. Moreover, there are no published normative data for recall consistency in depressed non-ECT samples that could be used to compare with depressed ECT samples. For example, what are the expected ranges for both baseline scores and percent recall consistency after 2 months for a 25-year-old man, after 6 months for a 70-year-old woman, or after 1 month for a depressed 55-year-old woman? Where can this information be found?

This means that even with the best intentions in the world, neither Dr Sackeim nor anybody else can actually know what is being measured or how well it is being measured and, in the absence of appropriate normative data, how to interpret findings beyond simply identifying differences between groups. The original developers of the CUAMI actually put this very well themselves when they appropriately acknowledged some of these limitations nearly 2 decades ago (page 513).
“The relatively small size of the normal control group \([n = 16]\) and the absence of retesting of this group were important limitations. Without normative information on the extent of inconsistency in recall over time, it is impossible to determine whether the treatment groups that showed the least retrograde amnesia after ECT, nonetheless, had short- or long-term deficits. Rather, the demonstration of differences among the treatment conditions can only be interpreted as indicating relative (as opposed to absolute) differences in the magnitude of retrograde amnesia.”

It is a pity that they did not take up this challenge, and it is a bit of a wonder that Dr. Sackeim now takes us to task like naughty schoolchildren for having the temerity to point out the same straightforward obvious limitations of these instruments. The CUAMI instruments mostly provide a measure of recall consistency that do not account for baseline performance (no matter how poor this might be). As previously noted, it is not possible to improve performance when compared to the baseline assessment. Indeed, it is possible for somebody with a terrible performance at baseline to have “better” percent consistency of recall than somebody with a normal rate of consistency loss! Because of this 2-stage design (ie, a measure of retrospective memory function at baseline followed by a different measure of recall consistency), it is not possible to derive an effect size for the CUAMI for an individual group.

Because of all the above, the CUAMI is not considered to be a validated neuropsychological measure. For example, it is not included in standard major neuropsychological reference texts as a measure for autobiographical memory. We were therefore unable to include it in a previous meta-analysis of the effects of ECT on cognitive performance as measured by reliable and validated neuropsychological tests. In a subsequent meta-analysis in which we allowed inclusion of nonvalidated tests, we did include studies using the CUAMI, although we did not extract effect sizes as incorrectly described by Dr. Sackeim. We clearly stated that effect sizes cannot be generated with these data that only allow comparisons between groups.

In an attempt to move forward, we have devised a new scoring system for the CUAMI-SF and have published reliability and validity data for both normal and depressed persons using this new approach. For the first time in the history of the CUAMI’s use, our validation study used a priori theoretically based hypotheses, integrating existing knowledge of autobiographical memory in general and of depression in particular, knowledge that Dr. Sackeim is consistently trying to gloss over. In this vein, we are continuing to work on clearing the water to measure and characterize the specific effect of ECT on autobiographical memory. We do so by using the new validated scoring system, which allows the distinction between retrograde amnesia associated with ECT and normal or mood-associated loss of consistency with the passage of time.

The biggest disservice to patients is to continue equating loss of autobiographical memory consistency to retrograde amnesia and to use them as synonymous (as Dr. Sackeim does throughout his paper) in the simplistic way it has been done in the past. During the past 20 years, what the CUAMI has been capturing after an ECT course is an indistinguishable mix of loss of consistency (due to the effect of time), lack of specificity (due to the effect of depression) and retrograde amnesia (due to the effect of ECT). This literature has not informed us on the nature, extent, and duration of retrograde amnesia attributable to ECT. It patently has not.

We would be delighted to see a peer-reviewed reliability and validity study of the CUAMI and CUAMI-SF. Indeed, it would be good scientific judgment, and not at all absurd, to do this before even contemplating filling the bath.

REFERENCES