

Psychometric Equivalence of Proprietary and Non-Proprietary Neuropsychological Test Configurations in the NACC Uniform Data Set



Michael Lee¹(mkeomec@uw.edu), Seo-Eun Choi¹, Phoebe Scollard², Connie Nakano¹, Kari A. Stephens³, Emily Trittschuh^{3,4}, Jesse Mez⁵, Andrew J. Saykin⁶, Timothy Hohman⁷, Paul K. Crane¹, Shubhabrata Mukherjee¹

1) Department of Medicine, University of Washington, 2) Université de Bordeaux, Bordeaux, France 3) Department of Psychiatry and Behavioral Sciences, University of Washington 4) VA Puget Sound Health Care System, GRECC 5) Department of Neurology, Boston University Medical Center 6) Department of Neurology and Radiology, Indiana University 7) Vanderbilt Memory and Alzheimer’s Center, Vanderbilt University Medical Center.

Background

- The National Alzheimer’s Coordinating Center (NACC) replaced proprietary neuropsychological instruments with non-proprietary alternatives with the transition from Uniform Data Set (UDS) version 2 to UDS version 3 in 2015.
- Traditional approaches to addressing battery transitions cannot assess preservation of domain-level measurement precision and clinical validity.
- We compare proprietary and non-proprietary test configurations using co-calibrated item bank parameters on two criteria: (i) measurement precision and (ii) association with clinical dementia severity in the NACC crosswalk sample.

Methods

- In 2014, NACC crosswalk study participants with a range of dementia status were administered both the UDS 2 and UDS 3 neuropsychological batteries (**Table 1**).
- We used previously established co-calibrated item parameters to derive harmonized scores for three cognitive domains: memory, executive functioning, and language across proprietary and non-proprietary sets of items
- The proprietary and non-proprietary items represent components that differ between UDS versions. Overlapping items were administered in both UDS1/2 and 3 batteries (**Table 2**).
- For each domain, we defined two configurations:
 - (a) Proprietary configuration: proprietary + overlapping UDS items
 - (b) Non-proprietary configuration: non-proprietary + overlapping UDS items
- We derived standard error of measurement (SEM) curves for each configuration across the range of each cognitive domain.
- We examined associations between domain scores and the Clinical Dementia Rating (CDR) Global score, dichotomized as no impairment (CDR = 0) versus dementia (CDR = 1-3), excluding participants with CDR = 0.5. We evaluated discriminative accuracy using areas under the ROC curve (AUC), and applying the DeLong test to compare the performance of the logistic regression models (**Figure 1**).

Table 1. Participant characteristics for the crosswalk sample by clinical diagnosis

	Normal cognition (n=510)	Impaired but not MCI (n=51)	Mild cognitive impairment (n=154)	Dementia (n=240)	Total (n=955)
Age at enrollment, mean (SD)	70.1 (8.8)	68.0 (9.3)	73.4 (8.5)	72.0 (9.5)	71.0 (9.1)
Age at crosswalk visit, mean (SD)	74.5 (9.3)	72.3 (9.5)	77.1 (9.3)	75.0 (10.3)	74.9 (9.6)
Female sex, n (%)	337 (66%)	20 (39%)	71 (46%)	110 (46%)	538 (56%)
Years of education, mean (SD)	16.2 (2.6)	16.3 (2.7)	15.5 (2.9)	15.6 (3.0)	15.9 (2.8)
CDR = 0	488	13	12	1	514
CDR = 0.5	22	38	139	77	276
CDR = 1-3	0	0	3	162	165

Table 2. Test item breakdown by cognitive domain: proprietary, non-proprietary and overlapping

	Memory		Executive Functioning		Language	
	UDS 2: Proprietary Items	UDS 3: Non-proprietary Items	UDS 2: Proprietary Items	UDS 3: Non-proprietary Items	UDS 2: Proprietary Items	UDS 3: Non-proprietary Items
Non-overlapping	Logical Memory (immediate recall)	Craft Story (immediate recall)	Digit Span (forward)	Number Span (forward)	Boston Naming Test	Multilingual Naming Test
	Logical Memory (delayed recall)	Craft Story (delayed recall)	Digit Span (backward)	Number Span (backward)		
Overlapping	none		Trail Making Test: A Trail Making Test: B		Animal Fluency Vegetable Fluency	

Results

- Measurement precision was plotted using SEM curves comparing proprietary and non-proprietary test configurations (**Figure 2**).
- The proprietary and non-proprietary configurations demonstrate similar AUC and sensitivity (**Figure 1**).
- Odds ratios of CDR and domain score association were not significantly different for executive functioning and language. Slight significant difference was detected for memory (p=0.048), but the effect sizes in proprietary and non-proprietary are very large and the difference between them is not clinically meaningful.

Figure 1. Area under the ROC comparison

The proprietary (blue, solid) and non-proprietary (red, solid) configurations across the cognitive domains of memory, executive functioning and language.

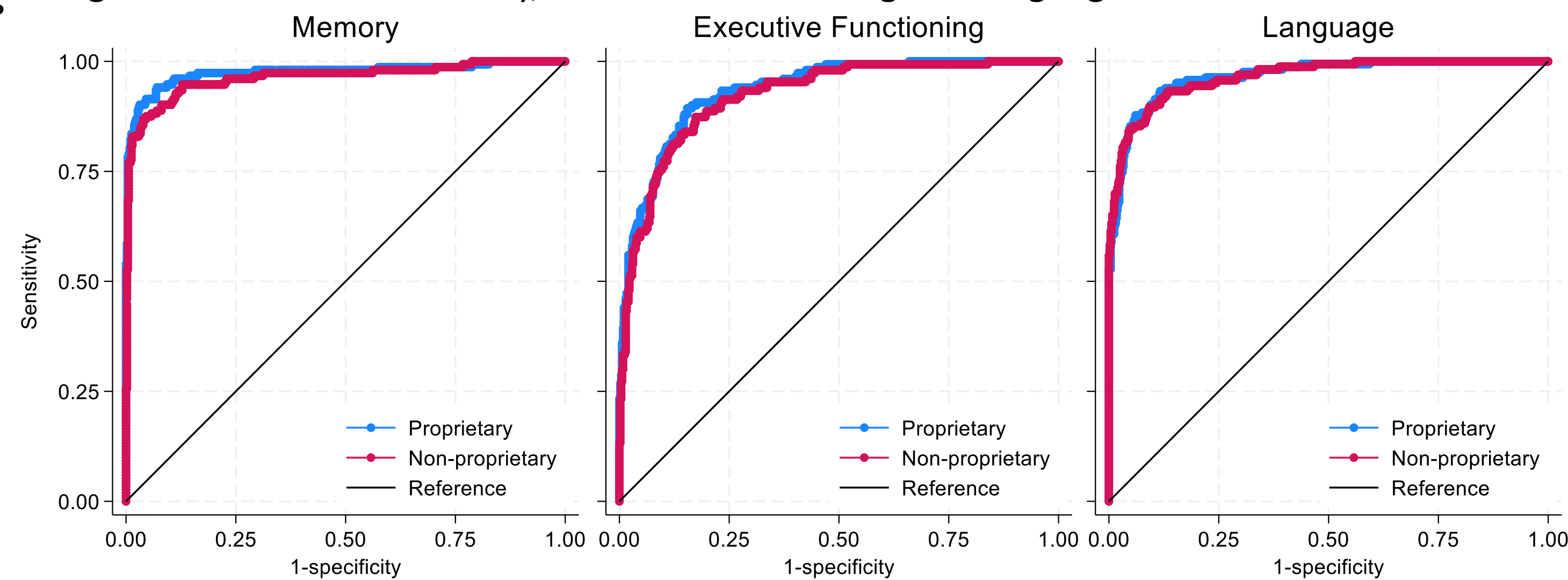
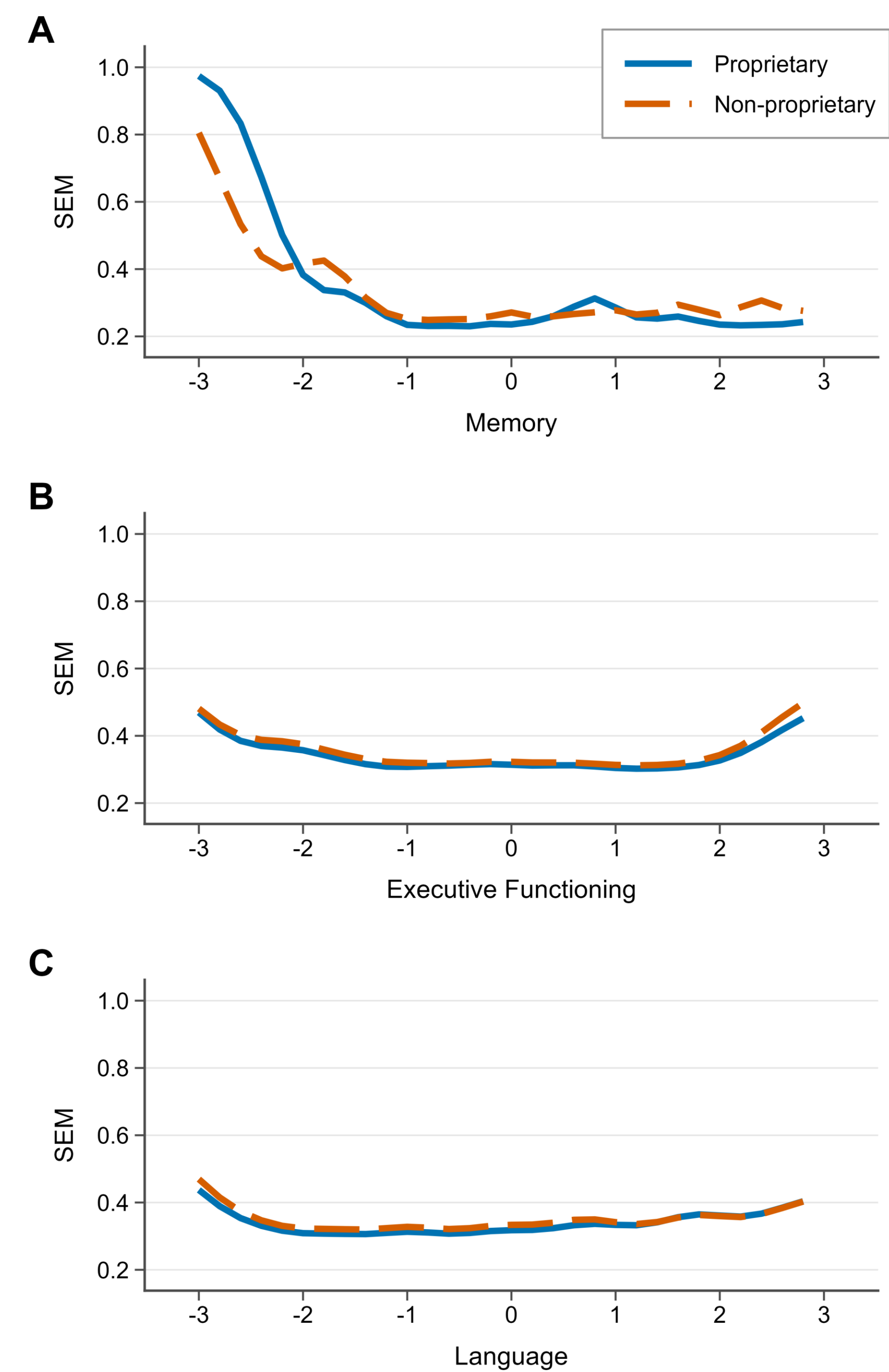


Figure 2. Standard error of measurement (SEM) curves

The proprietary (blue, solid) and non-proprietary (orange, dashed) configurations across the cognitive domains of memory (A), executive functioning (B) and language (C). Lower SEM values indicate greater measurement precision.



Discussion

- UDS3 non-proprietary neuropsychological instruments have psychometric properties similar to those of the proprietary instruments they replaced.
- We observed somewhat lower SEM for the proprietary memory configuration at the extreme low end of the latent trait distribution. However, this difference was confined to a region with relatively few participants and did not translate into meaningful differences in associations with clinical severity.
- The same approach can be applied to other research studies with battery changes due to licensing constraints, evolving measurement priorities or the need for cross-cultural adaptation.
- The approach can also be used to inform new studies designing batteries using non-proprietary instruments.