Relevant study aimed at providing normative data on the BAC-A in a Mandarin-speaking population in Taiwan. This work is of relevance for both clinical and research settings. It will allow health professionals to perform assessments of BD patients and be able to better define the severity of cognitive impairment in this population. The paper is well written and the data has been carefully analyzed. I particularly liked the attention paid to describing the linguistic validation and the author’s critical insight into weaknesses of the design. I have some comments related to the methodology and result section of this paper. I think they should be easy to address and I would encourage the authors to implement these changes in their revisions.

1. Although the BACS has been validated in Chinese patients with schizophrenia/ substance use disorders I was surprised to see that you worked only on the “new” items included in the BACA. Normative references should be generated based on the complete version of the BACA and not only the new items. Indeed, factors such as sequence of the tests, length of the assessment, motivation etc. play a role (especially in elderly population) and normative references of the same tests administered as part of the BACS may differ from those of the same tests administered as the BACA. I would suggest to do the following. You could either be very specific as to which tests you are going to include in your analyses from the very start (abstract, manuscript) or decide to analyze all the BAC-A tests.
2. If you do include all the BACA tests it’s important to describe all subscores in detail.
3. Given that you ran MANCOVAs I was surprised to see that your tables provided B scores and confidence intervals etc. why didn’t you provide F etc. ? Or was this a regression? Either way please make to describe exactly what kind of analyses you performed and please provide effect size of age-related differences in performance.
4. In terms of your multiple comparison correction. It’s extremely strict. Why is that? Did you consider running separate MANCOVAs with subset of similar variables or use FDR, sampling etc.
5. Although you keep referring to “motor speed” and “processing speed” there are no timed tasks (e.g. assessed with a stopwatch or computer) in the BACS/A. Which tasks are you referring to? This comment hints at the fact that one should be very specific when talking about “traditional neuropsychological domains” as they are broadly defined and they do not mean the same thing for everybody. For instance “executive function” is way too vague to really describe a specific ability. Please make sure to address this and try to be more precise when you refer to cognitive/neuropsychological domains in the abstract and manuscript.
6. I would like to know more about the way testing was conducted: e.g. questionnaires filled out before BACA, order of test assignment (random? Same for everyone), breaks/pauses, duration of the test, was the BACA administered via pen and paper or computer/tablet, evaluation of fatigue among older individuals, what did participant think of the test? Did they provide feedback? Testing times (morning/afternoon).
7. In table 1 please provide mean age for each group, measures of intelligence (e.g. WAIS, WRAT, vocabulary score), employment status, medication (e.g. painkillers etc. ), medical history etc. all these things should be tracked to better compare these results with other populations. If these estimates were not collected please mention this in your limitations etc.
8. Discussion: it would be helpful to compare these results to those of similar batteries (BACS for instance) administered to healthy populations. Please address this.
9. Please add a figure (e.g. line graph) summarizing the most relevant results. It would improve the understanding of your findings.

Minor

-I would make sure to talk about “linguistic validation” in both your abstract and to replace the header “adaptation of the Chinese…”. This is one of the strengths of your work and an important step towards psychometric validation and should be highlighted.

-replace “declined performance” with “impaired”, it would read a bit better.

-the number of references is surprisingly low, please consider adding relevant references if suitable

-why only ethnic Han Chinese (unfamiliar readers may wonder if there is a specific reason,e.g. genetic? Cultural? language)