Reviewer's report

**Title:** Refining Developmental Coordination Disorder subtyping with multivariate statistical methods

**Version:** 1  **Date:** 5 December 2011

**Reviewer:** Peter H Wilson

**Reviewer’s report:**

The aim of the paper was to explore two multivariate methods for classifying sub-types of DCD—random forests (RF) and partial least-squares discriminant analysis (PLS-DA). Using a sample of 63 children and 49 input variables, three main sub-types were indicated: ideomotor, visual-spatial/constructional, and mixed dyspraxia. While the methods used are impressive and worth applying to the DCD population, there are several major issues that warrant consideration before publication.

First, I am anxious about purely data-driven approaches in an area where theory is growing and increasingly informed by new advances in the cognitive neurosciences. I see little rationale for the choice of measures, nor any informed discussion about current hypotheses on motor control and learning in children DCD. Data reduction techniques will always generate a solution based on one algorithm or another. But, more important, is a sound theoretical justification for the input variables in the first place; this was not a strong feature in the current paper. As well, clustering can also be conceived at the level of motor control (e.g., force control, timing, online control, etc.), and not just as an ability-based scheme. These different possible conceptualisations need to be weighed up from the outset.

Second, I’m not sure that 63 children is a sufficient sample with 49 input variables, no matter how recent the analytic technique. There needs to be some strong a priori reasoning and justification for a sample size like this. Moreover, there is insufficient information on screening to reassure the reader about the diagnostic status of these children: Were they referred? What were there comorbid issues? Severity of DCD? Etc.

Third, the conceptual synthesis is cursory. It is not enough to simple re-state the merits of the analytic approach without considering whether the cluster solution can be externally validated. Again, this brings us back to the current theory and an informed choice of validation measures.

On balance, the paper needs substantial revision--it has merit, but is not yet a complete and persuasive manuscript.

**Level of interest:** An article of importance in its field

**Quality of written English:** Needs some language corrections before being
published

**Statistical review:** No, the manuscript does not need to be seen by a statistician.