

Comparative Performance of Video and Record Form Review Methods for Bayley Administration Fidelity

Whitney Wood, PhD¹, Laura Miller, PhD¹, Jess Segura, BSc¹, & Pam Ventola, PhD^{1,2}
¹ Cogstate, ² Yale Child Study Center

Methodological Issue Addressed

The relative sensitivity of record form review only vs. video and record form review when monitoring error identification in Bayley administrations in rare neurodevelopmental disorder (NDD) clinical trials.

Background

Ensuring the fidelity of Bayley administrations is essential to data integrity in rare NDD clinical trials. Administration errors such as inaccurate start or stop points, incorrect item delivery, and scoring inaccuracies can threaten endpoint validity.

While record-form review offers insight into scoring and documentation accuracy, it cannot evaluate examiner behavior. In contrast, video review allows full observation of administration procedures in addition to review of the record form but requires greater resources.

Understanding the relative sensitivity of these approaches is essential for optimizing central-review strategies and ensuring high-quality data.

Methods

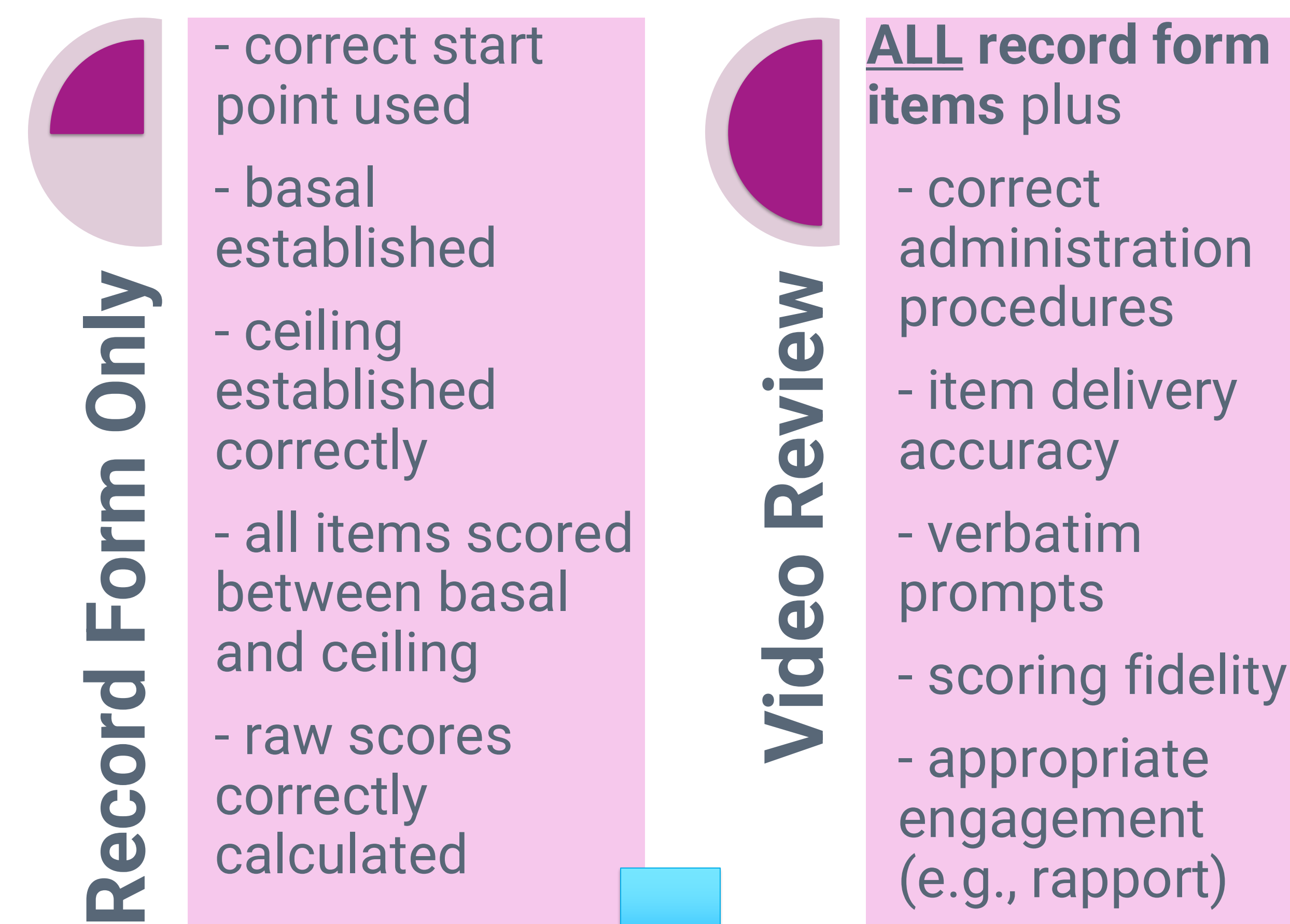
Bayley-III and Bayley-4 administrations (N = 520) from several ongoing and completed NDD trials were independently evaluated using one of the two approaches:

- Record-Form Only Review (n = 408)
- Video Review (n = 112)

Each administration was reviewed using *either* record form review or video review.

Methods

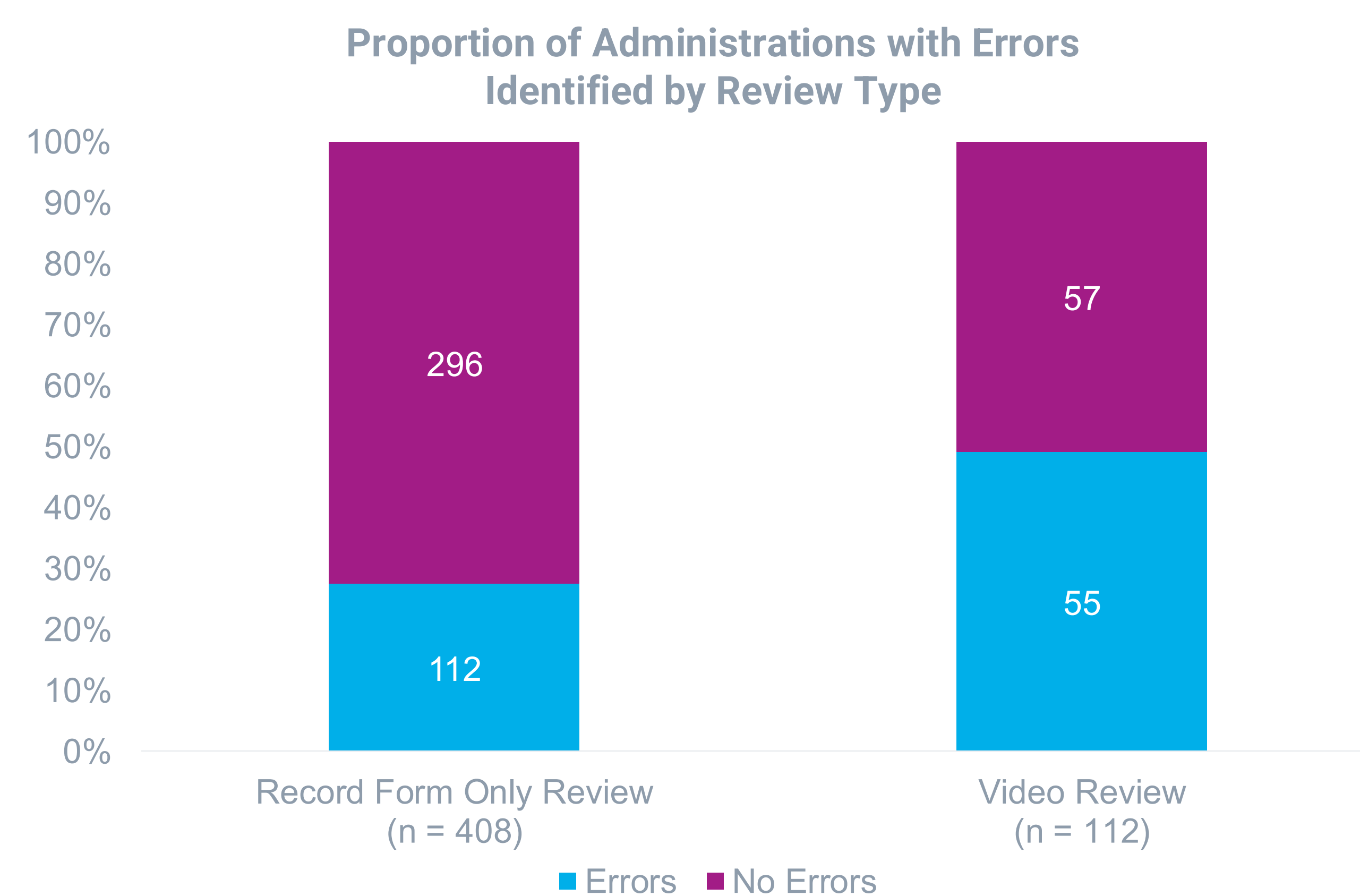
Trained central reviewers evaluated the following:



Administrations were classified dichotomously as having (a) errors or (b) no errors.

Results

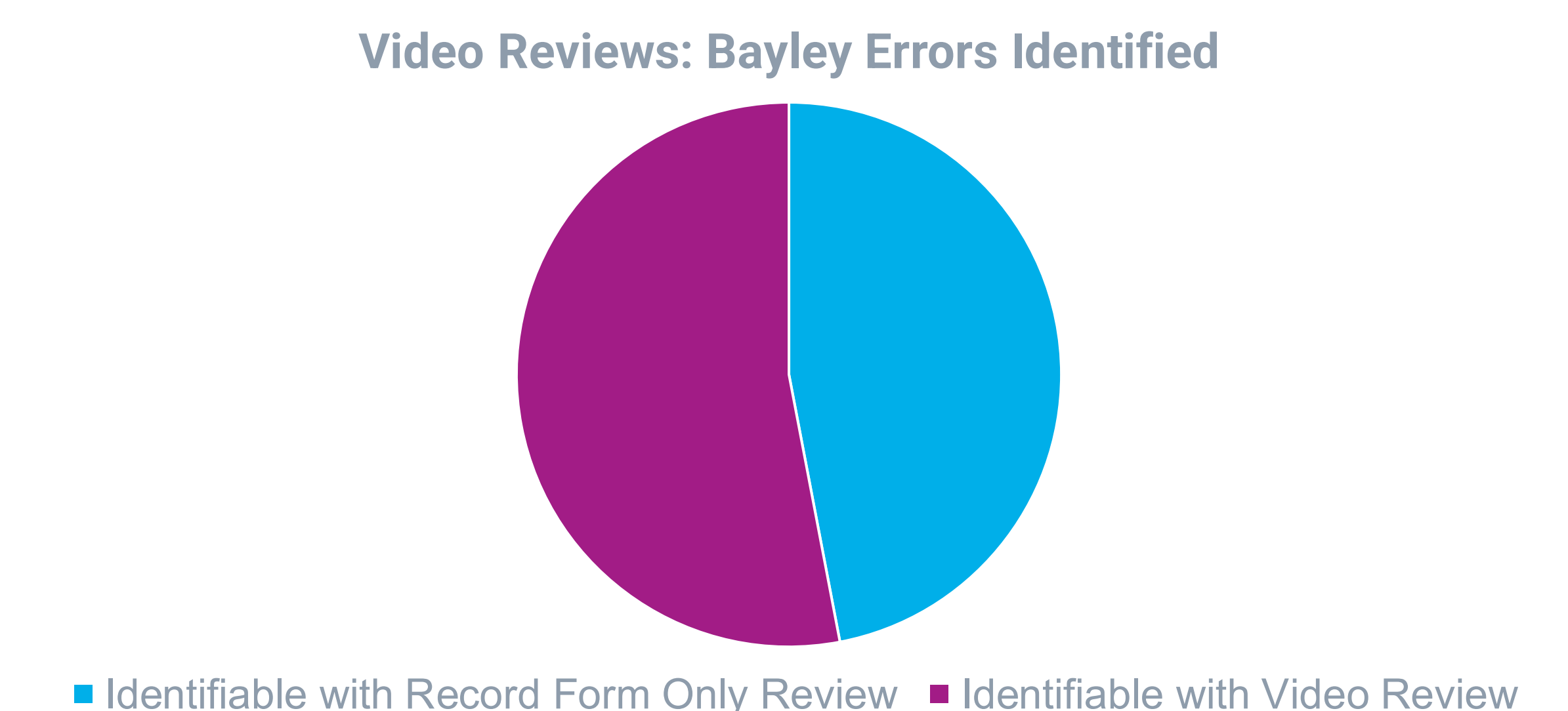
Significant association between review method and error detection, $\chi^2(1) = 18.90, p = .000014$.



Video review identified a substantially higher proportion of administrations with errors (49%) compared with record-form review (27%).

Results

Of the 55 video review administrations with errors, 27 (47%) contained errors identifiable through record-form review *only*, while 28 (53%) contained performance-based errors that were detected only through video review (e.g., incorrect item delivery, scoring fidelity, verbatim prompts).



Conclusion

1. Video review of Bayley administrations may offer greater sensitivity for monitoring administration quality in NDD trials.
2. Record-form review of the Bayley remains valuable for identifying scoring inaccuracies, but reliance on record forms alone may underestimate administration error rates.

References

Bayley, N. (2005). Bayley Scales of Infant and Toddler Development (3rd ed.). Pearson.

Bayley, N., & Aylward, G. P. (2019). Bayley Scales of Infant and Toddler Development (4th ed.). Pearson.