An examination of measurement procedures and baseline behavioral outcomes in single-case research

James E. Pustejovsky (pusto@austin.utexas.edu), Daniel M. Swan (dswan@utexas.edu), & Kyle English (kenglish@utexas.edu)

Overview

There has been growing interest in using statistical methods to analyze data and estimate effect size indices from studies that use single-case designs (SCDs), as a complement to traditional visual inspection methods. The validity of a statistical method rests on whether its assumptions are plausible representations of the process by which the data were collected, yet there is evidence that some assumptions—particularly regarding normality of error distributions—may be inappropriate for single-case data. To develop more appropriate modelling assumptions and statistical methods, researchers must attend to the features of real SCD data.

In this study, we examine several features of SCDs with behavioral outcome measures in order to inform development of statistical methods. Drawing on a corpus of seven systematic reviews covering a range of intervention classes and outcome constructs, which includes over 300 studies and approximately 1800 cases, we report the distribution of study designs, distribution of outcome measurement procedures, and features of baseline outcome data distributions for the most common types of measurements used in single-case research. We discuss implications for the development of more realistic assumptions regarding outcome distributions in SCD studies, as well as the design of Monte Carlo simulation studies evaluating the performance of statistical analysis techniques for SCED data.

Characteristics of included systematic reviews

| Review	Authors	Population	Interventions	Outcomes	Studies
Choice-making	Shogren, Faggella-Luby, Bae, & Wehmeyer (2004)	Individuals with disabilities	Choice-making opportunities	Disruptive behavior, on-task behavior, social skills	13
FBA	Gage, Lewis, & Stichter (2012)	Students with/at risk for emotional/ behavioral disorders	Functional behavioral assessment	Problem behavior	67
Group contingencies	Maggin, Pustejovsky, & Johnson (2017)	Students with challenging behavior	School-based group contingencies	Social interaction, academic engagement, disruptive behavior	40
Object play	Barton, Sweeney, & Gossett (2016)	Young children with disabilities	Least-to-most prompting, positive reinforcement	Object play	11
Peer management	Dart, Collins, Klingbeil, & McKinley (2014)	Students	Peer management interventions in school settings	Disruptive behavior, on-task behavior, social skills	29
PRT	Verschuur, Didden, Lang, Sigafoos, & Huskens (2014)	Individuals with ASD	Pivotal response training	Self-initiations, communication/language	31
Social skills	Ledford, King, Harbin, & Zimmerman (2016)	Individuals with ASD	Antecedent social skills	Pro-social behaviors	112

Study characteristics by review

Review	Participants per study			Series per participant			Baseline sessions per series		
	Mean	Median	IQR	Mean	Median	IQR	Mean	Median	IQR
Overall	3.1	3	2-4	2.0	1	1-2	11.1	7.0	5-12
Choice-making	2.5	3	1-3	2.3	2	1-3	7.4	5.5	4-9
FBA	2.1	2	1-3	1.5	1	1-2	8.6	7.0	4-10
Group contingencies	3.3	3	2-4	1.5	1	1-2	7.4	5.0	4-9
Object play	3.8	3	3-4	1.3	1	1-2	11.3	9.5	5-12
Peer management	3.0	3	2-3	1.6	1	1-2	10.8	8.0	5-13
PRT	4.4	3	3-5	3.1	2	1-5	9.5	6.0	4-11
Social skills	3.2	3	3-4	2.0	2	1-2	13.9	9.0	6-16

Distribution of initial baseline phase length across seven systematic reviews of SCDs
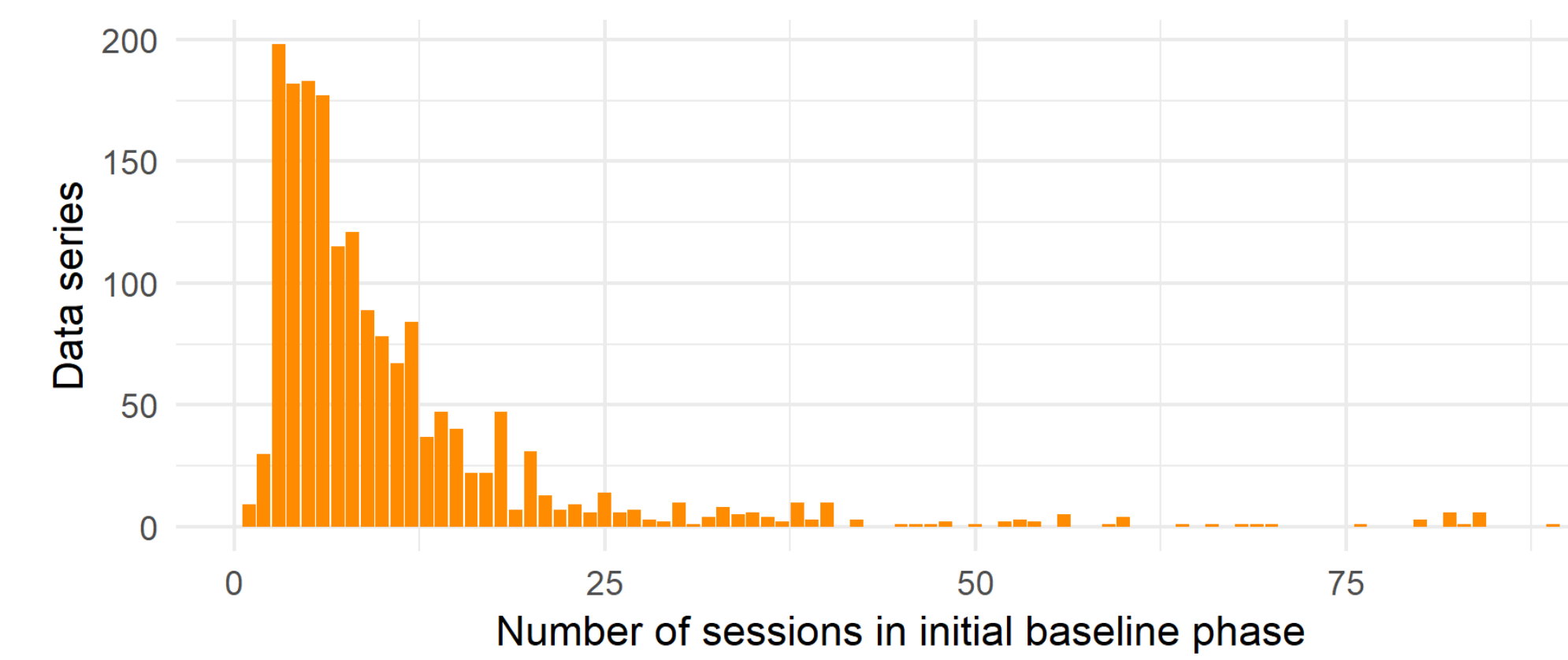
Event counting outcomes: mean baseline levels and mean-variance relationships

Scatterplots of variance versus sample mean baseline frequency for event count outcomes, with marginal distributions of mean baseline frequency. Each point represents one data series, with size corresponding to baseline phase length. Dashed lines represent unit slopes, where variance is equal to mean. Blue curves depict local linear regressions of variance as a function of mean. Dots along the horizontal axis indicate the quartiles of the distribution.
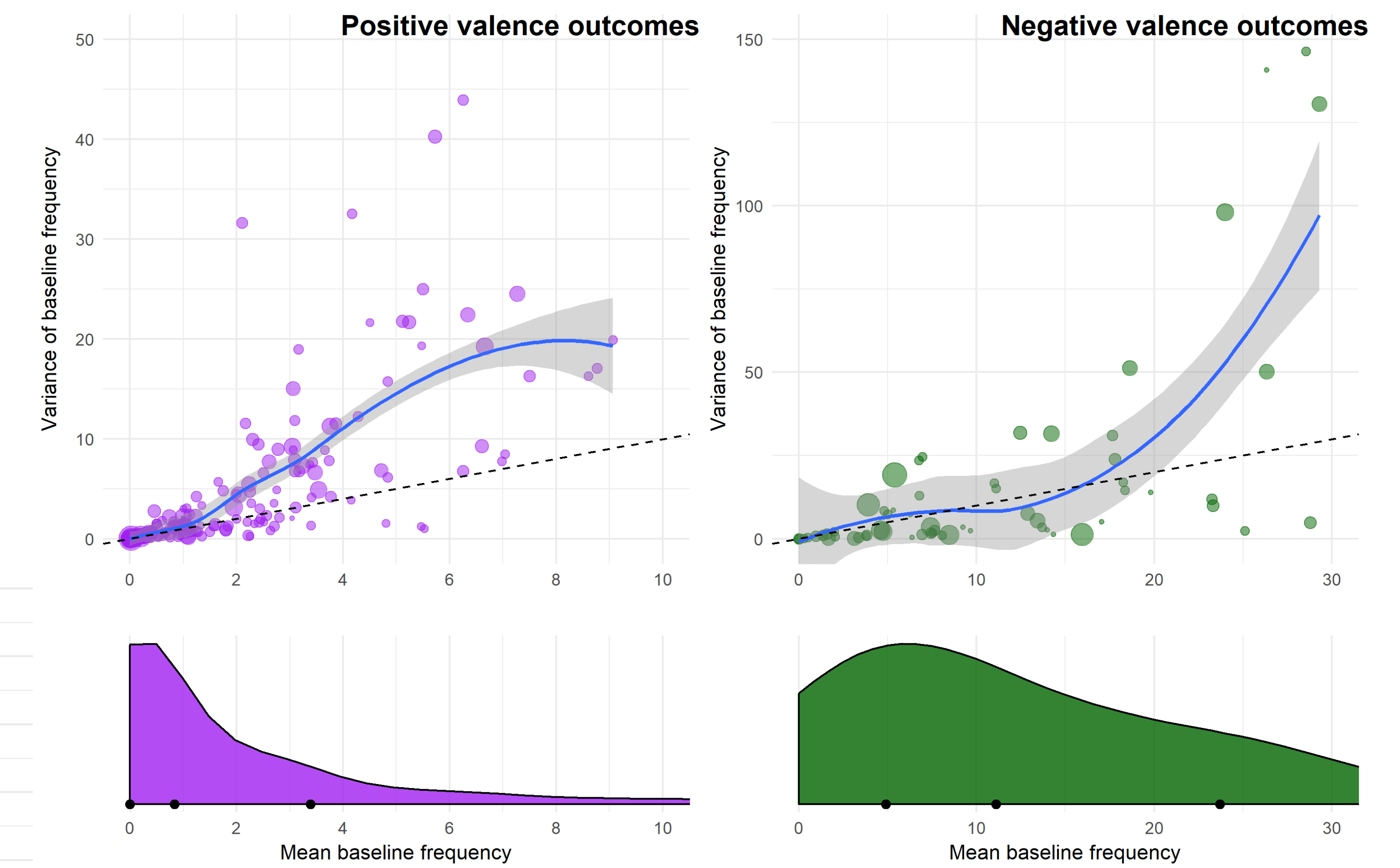
## Distribution of study designs by review

Review	Multiple baseline across participants	Treatment reversal	Multiple baseline across behaviors/ settings	Multiple probe across participants	Other
Overall	145 (48%)	102 (34%)	31 (10%)	13 (4%)	12 (4%)
Choice-making	2 (15%)	10 (77%)	1 (8%)	-	-
FBA	15 (22%)	39 (58%)	9 (13%)	-	4 (6%)
Group contingencies	11 (28%)	28 (70%)	1 (2%)	-	-
Object play	8 (73%)	-	-	3 (27%)	-
Peer management	15 (52%)	11 (38%)	3 (10%)	-	-
PRT	27 (97%)	3 (10%)	-	-	1 (3%)
Social skills	67 (60%)	11 (10%)	17 (15%)	10 (9%)	7 (6%)

Measurement procedures by review

| Procedure	Overall	Choice-making	FBA	Group contingencies	Object play	Peer management	PRT	Social skills
Partial interval recording	650 (36%)	22 (31%)	135 (67%)	52 (26%)	26 (48%)	64 (44%)	247 (58%)	104 (14%)
Event counting	419 (23%)	7 (10%)	13 (6%)	61 (31%)	21 (39%)	42 (29%)	61 (14%)	214 (30%)
Success (fixed)	363 (20%)	-	-	3 (2%)	3 (6%)	8 (6%)	72 (17%)	277 (38%)
Success (variable)	97 (5%)	11 (15%)	2 (1%)	-	-	4 (3%)	28 (7%)	52 (7%)
Momentary time sampling	86 (5%)	1 (1%)	16 (8%)	66 (34%)	-	3 (2%)	-	-
Continuous recording	80 (4%)	6 (8%)	16 (8%)	15 (8%)	-	3 (2%)	-	40 (6%)
Whole interval recording	69 (4%)	10 (14%)	19 (9%)	-	-	15 (10%)	3 (1%)	22 (3%)
Response latency	19 (1%)	-	-	-	-	-	6 (1%)	13 (2%)
Task check-list	18 (1%)	10 (14%)	-	-	-	5 (3%)	-	3 (0%)
Rating scale	17 (1%)	4 (6%)	1 (0%)	-	4 (7%)	-	8 (2%)	-

Distribution of session lengths in data series measuring free-operant behaviors

Dots indicate the quartiles of the distribution for each review. Session lengths over 30 minutes are not depicted.
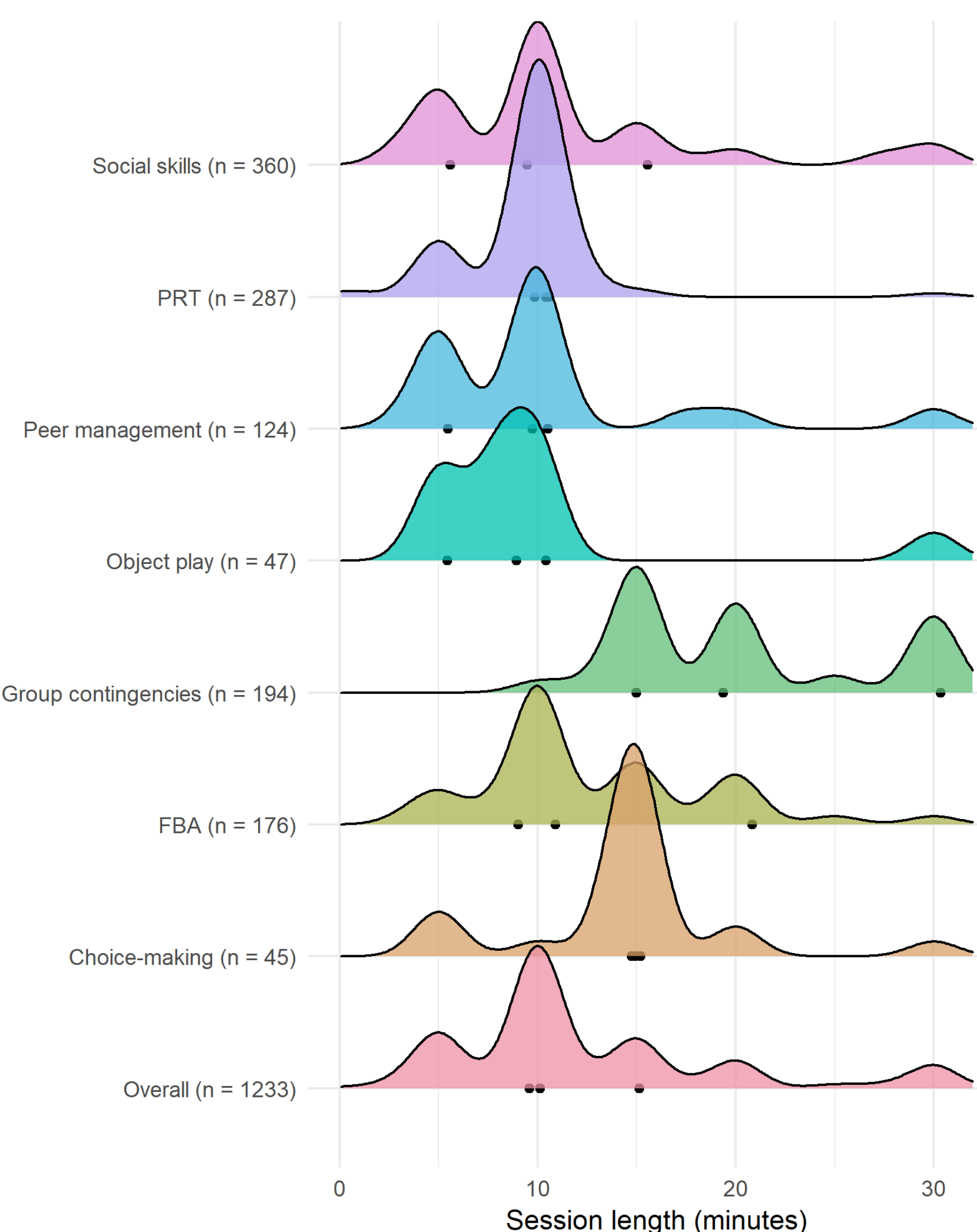
Proportion outcomes: mean baseline levels

Distributions of mean baseline proportions for partial interval recording (PIR), momentary time sampling (MTS), continuous recording (CR), and whole interval recording (WIR) data series. Left-hand plots depict positive-valence outcomes. Right-hand plots depict negative-valence outcomes. Dots along the horizontal axis indicate the quartiles of the distribution.
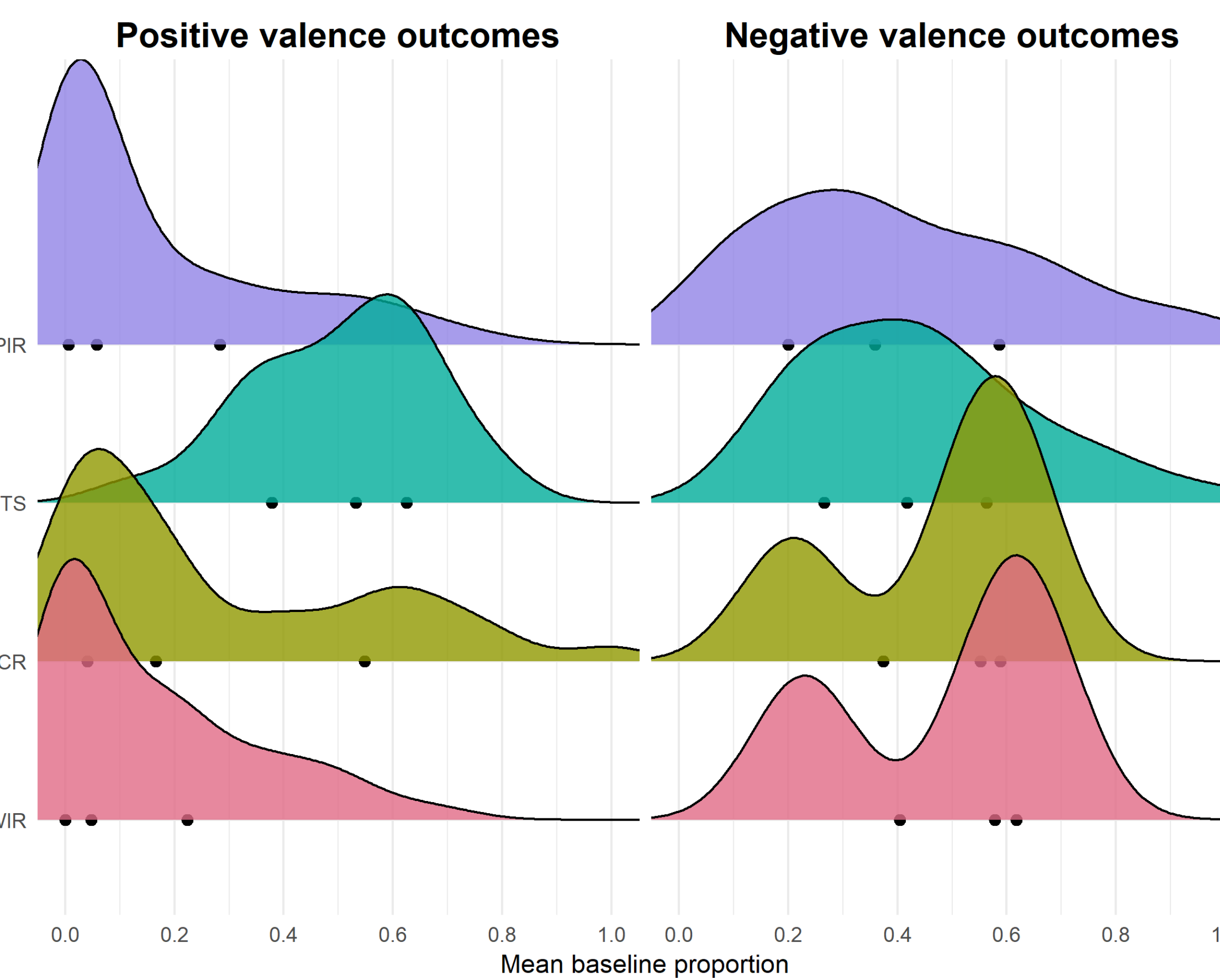
Proportion outcomes: mean-variance relationships

Scatterplots of sample variance versus sample mean baseline proportions for partial interval recording, momentary time sampling, continuous recording, and whole interval recording data series. Each point represents one data series, with size corresponding to baseline phase length. Curves depict local linear regressions of variance as a function of mean.