

ORIGINAL ARTICLE

How Many Sleep Diary Entries Are Needed to Reliably Estimate Adolescent Sleep?

Michelle A. Short, PhD^{1,2}; Teresa Arora, PhD³⁻⁵; Michael Gradisar, PhD²; Shahrad Taheri, PhD, MBBS, FRCP⁴⁻⁶; Mary A. Carskadon, PhD^{1,7}

¹Zayed University, College of Natural and Health Sciences, Department of Psychology, Abu Dhabi, United Arab Emirates; ²School of Psychology, Flinders University, Bedford Park, South Australia, Australia; ³Department of Psychology, College of Natural and Health Sciences, Zayed University, Abu Dhabi, United Arab Emirates; ⁴Department of Medicine, Weill Cornell Medicine, Doha, Qatar; ⁵Department of Medicine, Weill Cornell Medicine, New York, NY; ⁶Department of Medicine, King's College London, London, UK; ⁷E.P. Bradley Hospital Sleep Research Laboratory, Providence, RI

Study Objectives: To investigate (1) how many nights of sleep diary entries are required for reliable estimates of five sleep-related outcomes (bedtime, wake time, sleep onset latency [SOL], sleep duration, and wake after sleep onset [WASO]) and (2) the test–retest reliability of sleep diary estimates of school night sleep across 12 weeks.

Methods: Data were drawn from four adolescent samples (Australia [$n = 385$], Qatar [$n = 245$], United Kingdom [$n = 770$], and United States [$n = 366$]), who provided 1766 eligible sleep diary weeks for reliability analyses. We performed reliability analyses for each cohort using complete data (7 days), one to five school nights, and one to two weekend nights. We also performed test–retest reliability analyses on 12-week sleep diary data available from a subgroup of 55 US adolescents.

Results: Intraclass correlation coefficients for bedtime, SOL, and sleep duration indicated good-to-excellent reliability from five weekday nights of sleep diary entries across all adolescent cohorts. Four school nights was sufficient for wake times in the Australian and UK samples, but not the US or Qatari samples. Only Australian adolescents showed good reliability for two weekend nights of bedtime reports; estimates of SOL were adequate for UK adolescents based on two weekend nights. WASO was not reliably estimated using 1 week of sleep diaries. We observed excellent test–retest reliability across 12 weeks of sleep diary data in a subsample of US adolescents.

Conclusion: We recommend at least five weekday nights of sleep diary entries to be made when studying adolescent bedtimes, SOL, and sleep duration. Adolescent sleep patterns were stable across 12 consecutive school weeks.

Keywords: sleep diary, adolescent, sleep, reliability, test–retest.

Statement of Significance

Reliable estimates of sleep are crucial to clinicians and researchers who seek to quantify the typical sleep of an individual or group. The absence of reliable estimates means that any associations between sleep and other constructs (cognitive functioning, health outcomes, etc.), or observations of change in sleep patterns over time (in the case of longitudinal or treatment studies), become meaningless. This study is the first to determine how many nights of sleep diary estimates are needed to reliably estimate sleep in adolescents, as well as the test–retest reliability of consecutive weeks of sleep diary entries. Future research is needed to address this critical gap regarding younger populations.

INTRODUCTION

Sleep diaries are widely used in sleep research and clinical practice and are considered the gold standard for subjective sleep assessment.¹ They have the advantage of being easy to use, portable, and low cost. As sleep diaries measure sleep prospectively, many of the limitations of retrospective self-reported sleep are avoided, including recency effects (individuals' tendency to recall recently occurring examples) and recall biases. Despite their widespread use, limited research has evaluated how many nights of sleep diary records are needed for reliable sleep estimates in adults,² and none is available in adolescents. The lack of such research frames an important gap in the evidence for accurate estimates of typical sleep from daily diaries. Aggregation of multiple nights of sleep estimates reduces the effect of intraindividual night-to-night variability and provides greater accuracy of sleep estimates.

A number of factors affect the number of sleep diary nights needed for reliable sleep pattern estimates. These factors include the aspect of sleep assessed (ie, sleep onset latency [SOL], total sleep time, bedtime, wake time, etc.), the period of sleep aggregated (ie, school night sleep, weekend sleep, weekly sleep, etc.), and the sample population (ie, healthy school-attending teenagers, those with clinical sleep disorders [ie, delayed sleep–wake phase disorder, insomnia], etc.).³ Common among these factors

is what drives differences in how many nights are needed for reliability, namely, intraindividual stability.² Greater stability requires the aggregation of more data points to provide accurate estimates of “typical” sleep. For example, in healthy older adults, Wohlgemuth and colleagues estimated that 1 week of sleep diary records were needed to provide adequate stability.² Among older individuals with insomnia, however, 2 weeks were needed to estimate SOL, while reliable estimates of wake after sleep onset (WASO) required 3 weeks of sleep diary records.² This difference was due to the increased variability in night-to-night SOL and WASO among the clinical group.

Adolescence is characterized by a high degree of variability in sleep,⁴⁻⁶ with systematic differences in sleep parameters occurring according to when the measurement is made and which variable is being measured. For example, school night sleep typically involves earlier bedtimes, earlier wake times, and a shorter sleep duration than weekend sleep.^{7,8} Due to the high degree of variability among adolescents, estimates of reliability derived from adult samples cannot be generalized to this age group. Thus, it is important to estimate reliability specifically for adolescents. Indeed, even within adolescent groups, known cross-cultural differences in sleep and sleep variability may affect reliability estimates.⁹⁻¹¹ Accordingly, the present study

examines how many nights of sleep diary records are needed to estimate SOL, bedtime, wake time, sleep duration, and WASO in culturally diverse groups of adolescents from Australia, Qatar, the United States, and the United Kingdom.

The present study also evaluates the test–retest reliability of sleep diary estimates of sleep to determine the stability of sleep estimates over several months. Longitudinal and clinical research that uses sleep as an outcome variable rely on the assumption that changes to sleep in the short-to-medium term are driven by changes in the independent variable (such as treatment), rather than by the passage of time or by measurement error. For example, a common test for clinical significance is the Reliable Change Index (RCI) statistic to determine whether or not a change in an individual's scores pre- and post-treatment is significant (ie, an $RCI > 1.96$).¹² The RCI can be applied to experimental case studies or to provide an indication of the percentage of individuals in a sample show clinical improvement. A common use of RCIs is to assess treatment effects in mental health disorders.^{13–15} Despite the frequent use of sleep diaries in treatment outcome research, RCIs for sleep diaries cannot be calculated because the sleep research field has not determined their test–retest reliability, which is needed within the RCI formula. To address this gap, 12 weeks of sleep diary data from a subsample of 55 US adolescents were examined to determine the test–retest reliability of sleep diaries.

METHODS

Australian Adolescents

Ethical approval for the study was granted by Flinders University Social and Behavioural Research Ethics Committee and the Department of Education and Children's Services. Participants were 385 adolescents (60% male) recruited from eight South Australian high schools and were aged 13–18 years ($M = 15.6$, $SD = 0.95$). All students attending these mainstream high schools were eligible to participate and there were no exclusion criteria. Informed consent was obtained from the Principal of each school, the participant, and a parent or guardian. The response rate was 84%.

Participants completed a pencil and paper sleep diary across 1 week during the school term. Variables included bedtime (“Went to bed last night at”), SOL (“Minutes until fell asleep”), sleep duration (“Slept this much last night”), WASO (“After you fell asleep, what is the total amount of time you spent awake during any night time awakenings?”), and wake time (“Finally woke at”). For each item, adolescents were instructed to answer with a specific time (eg, 9:35 pm) or duration (eg, 12 minutes). Participants also concurrently wore wrist actigraphy (Mini-motionlogger; Ambulatory Monitoring Inc. (AMI), Ardsley, NY) and called a time-stamped answering machine before bedtime and after they woke each day.

US Adolescents

US School Sample

A total of 311 students (35% male) were recruited from five Rhode Island high schools between late September and early December. Students were in grades 9–12 and ranged in age from 13 to 19 years ($M = 16.0$, $SD = 1.2$). Inclusion criteria were only the ability to understand and willingness to comply with the study's requirements.

US Subsample

Participants were 55 adolescents (40% male), aged 14.5–19.3 years, who attended high school in Rhode Island. Participants were recruited from the 311 adolescents from the previously described study.¹⁶ Exclusion criteria included diagnoses of sleep apnoea, narcolepsy, restless legs, major depressive disorder, bipolar disorder, schizophrenia, current eating disorder, or a major chronic medical condition, head trauma with a loss of consciousness, repeated involvement with the juvenile court system or current use of psychoactive substances known to affect the sleep/wake cycle or daytime alertness. Participants completed pencil and paper sleep diaries prospectively across one school semester from late January to mid-May. Only school weeks were included in the analyses.

The Lifespan Institutional Review Board (IRB) approved the two US studies, and informed consent was obtained from parents of participants under age 18 years, or from students aged 18 years or older. Pencil and paper sleep diary entries for both studies included bedtime (“Attempted to fall asleep at”), sleep duration (“Slept this much last night”), and wake time (“Finally woke at”). SOL (“Minutes until fell asleep”) entries were only available for the US subsample. The first week of sleep diary entries were used to analyze night-to-night stability for each group separately. Twelve weeks of school night estimates were included in the test–retest analyses. Participants also wore wrist actigraphy (Mini-motionlogger; AMI) and called a time-stamped telephone answering machine before bedtime and after rise time. Sleep diaries were collected weekly from participants.

Qatari Adolescents

Approval for Qatar's Ultimate Education for Sleep in Teenagers (QUEST) study was obtained from Qatar's Supreme Council of Education. Ethical approval for the study was granted by Weill Cornell Medicine in Qatar Institutional Review Board (14-00104). Head teachers at two international high schools located in Doha, Qatar, were approached and agreed to participate in the study. Parents of all students registered in years 7–11 from the two high schools received information about the study and opt out consent was provided. Willing students provided written assent. Students were excluded from study participation if they had a diagnosed sleep disorder(s), had travelled to a different time zone 4 weeks before providing sleep data, or took medication that affected their sleep (over-the-counter or prescribed). A total of 611 adolescents were recruited to the study (aged 10–18 years; 44% male). The main aim of this cross-sectional study was to identify sleep patterns in adolescents residing in Qatar. Participants were provided with wrist actigraphy concurrent with a pencil and paper 7-day sleep diary for completion during term time. A trained researcher explained to each participant, in a private room setting, how to complete the sleep diary. The sleep diary requested the following information: bedtime (time you went to bed and attempted to sleep), sleep time (time you think you fell to sleep), and wake time (time you finally woke up). From these variables, we calculated: sleep duration (difference between sleep onset and wake time in minutes) and SOL (difference between bedtime and sleep time in minutes). For the purposes of our analyses, we investigated number of school nights (one to five) and the number of weekend nights

(one to two) for bedtime, SOL, sleep duration, and wake time. A total of 245 adolescents returned their sleep diary with at least one entry.

UK Adolescents

Ethical approval for the Midlands Adolescent School Sleep Education Study (MASSES) was granted by the University of Birmingham Research Ethics Committee (ERN_08-437). In brief, eight secondary schools located in the Midlands region of the United Kingdom participated in the 3-year cohort study. Parents of students registered in years 7–9 at the schools (aged 11–13 years at baseline) received information about the study and opt in/out consent was provided (response rate 80%), according to the Head Teacher’s recommendation. Students were excluded from study participation if they had a diagnosed sleep disorder(s), had travelled to a different time zone 4 weeks before providing sleep data (at any of the three assessments), or took medication that affected their sleep (over-the-counter or prescribed). Willing students provided personal assent. A total of 892 adolescents were eligible to participate in the study (42.4% male). The main aim of the study was to assess changes in sleep patterns in UK adolescents over a 3-year period with annual sleep assessments (7 consecutive days/nights of wrist actigraphy and a corresponding sleep diary) during the same term each year. A trained researcher provided instructions for sleep diary completion to student volunteers in a group setting and participants were requested to return the sleep diary 1 week later. The pencil and paper sleep diary used in MASSES was the same as the one used in QUEST and the same sleep variables were calculated, as previously highlighted. A total of 293 participants returned the sleep diary in year 1; 261 returned the diary in year 2; and 216 returned the diary in year 3 of the study. To be included, sleep diaries required at least one entry

($n = 770$ across all three time points assessed from adolescents aged 11–16 years over the entire study period).

Statistical Analyses

All statistical analyses were performed using IBM SPSS Statistics Version 24 (IBM Corp., Armonk, NY). Table 1 includes the items used to measure each sleep variable across the four countries. The nightly means and SDs were calculated for each sleep variable according to country and are presented in Figure 1 and Table 2. A random effects analysis of variance was performed to determine the absolute agreement between sleep diary data across a specified number of nights. Intraclass correlation coefficients (ICCs) between .60 and .74 indicate good reliability, while ICCs above .75 indicate excellent reliability.¹⁷ These values were used as the cut point in our analyses. The average ICC was obtained for all seven nights, for each of the four samples and for each of the sleep parameters assessed (bedtime, SOL, sleep duration, wake time). Estimates of WASO were only available for the Australian cohort. The same process was repeated for five school nights (Saturday to Wednesday for Qatar; Sunday to Thursday for Australia, United Kingdom, and United States), and the analyses were repeated by removing one night at a time leaving four schools nights (Saturday to Tuesday for Qatar; Sunday to Wednesday for Australia, United Kingdom, and United States), then three school nights (Saturday to Monday for Qatar; Sunday to Tuesday for Australia, United Kingdom, and United States), and so on. The same process was repeated for weekend nights (Friday and Saturday for Australia, United Kingdom, and United States; Thursday and Friday for Qatar) and the ICCs were obtained. The ICCs for a single night were obtained from the analyses of the model where complete weekday (five school nights) or weekend (two weekend nights) data were entered.

Table 1—Items Used to Operationalize Sleep Variables in Each of the Sleep Diaries.

Bedtime	
United States	Attempted to fall asleep at _____ am pm (circle am or pm)
Australia	Went to bed last night at _____ am pm (circle am or pm)
Qatar and United Kingdom	Time you went to bed _____ (example entry of 8:00 pm was given in first row)
Sleep onset latency	
United States and Australia	Minutes until fell asleep _____
Qatar and United Kingdom	Researcher calculated variable as difference in minutes between “Time you went to bed” and “Time you think you fell to sleep”
Wake time	
United States and Australia	Finally woke at _____ am pm (circle am or pm)
Qatar and United Kingdom	Time you woke up _____ (example entry of 6:00 am was given in first row)
Sleep duration	
United States	Slept this much last night _____ hours and _____ minutes
Australia	Went to bed last night at _____ am pm (circle am or pm)
Qatar and United Kingdom	Total sleep duration (hours and minutes) _____ (example entry of 9 h 30 min was given in the first row)

Units of measurement were clock times for bedtime and wake time, minutes for sleep onset latency, and hours and minutes for sleep duration.

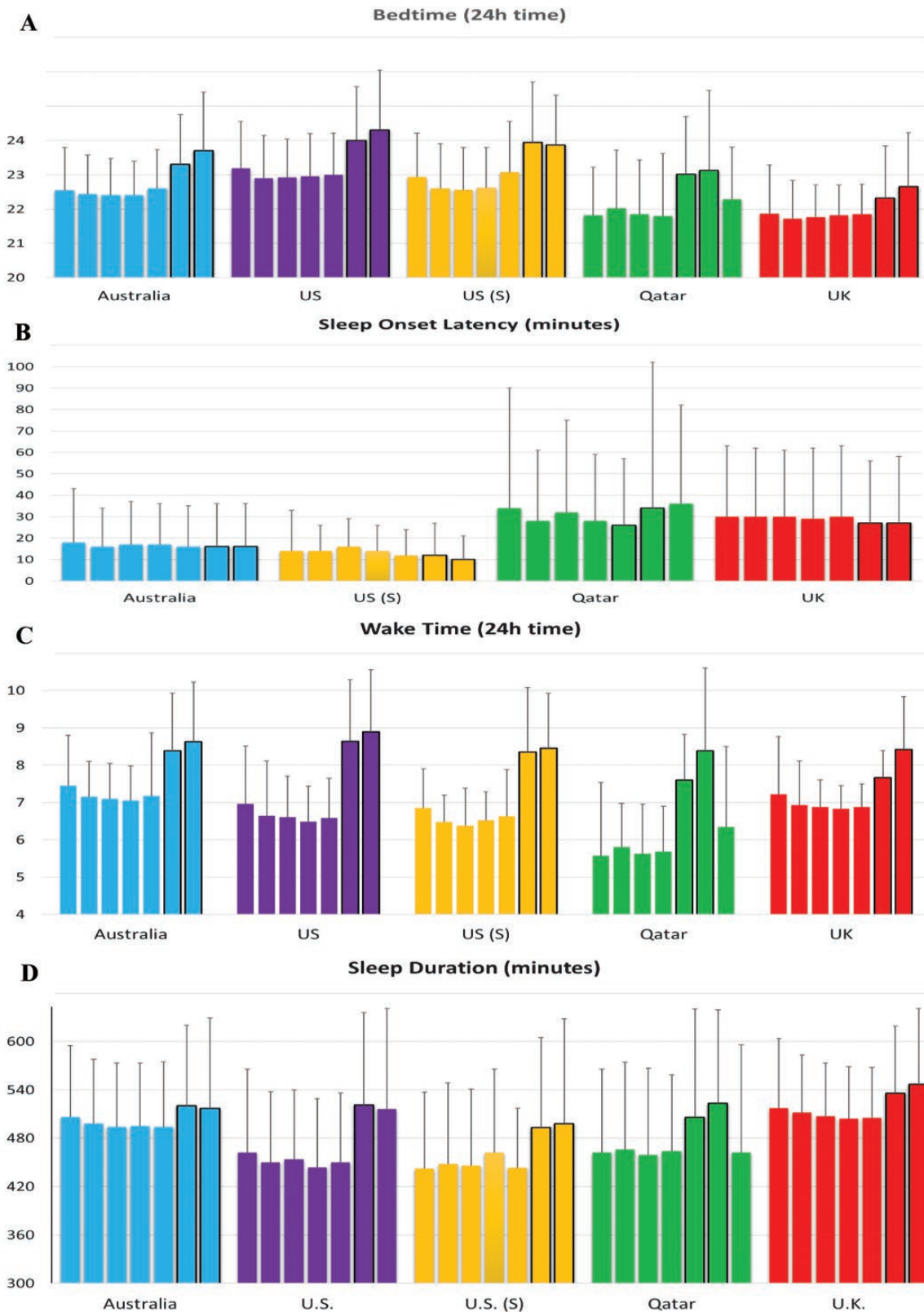


Figure 1—Mean (SD) values for bedtime (A), sleep onset latency (B), wake time (C), and sleep duration (D). Nights are Sunday night to Saturday night. Weekend nights are outlined in black.

Test-retest reliability was calculated for school night estimates of all sleep measures using 12 weeks of consecutive school night data from the US adolescents. School night

estimates were chosen as adolescents often present clinically due to problematic school night sleep, frequently due to the conflict in timing between their own sleep timing and that of

Table 2—Means (SD) for Consecutive School Nights and Weekend Nights.

	School night					Weekend night	
	1 st	2 nd	3 rd	4 th	5 th	1 st	2 nd
Bedtime							
Australia	22.6 (1.3)	22.4 (1.1)	22.4 (1.2)	22.4 (1.0)	22.6 (1.1)	23.3 (1.5)	23.7 (1.7)
United States	23.2 (1.4)	22.9 (1.3)	22.9 (1.1)	23.0 (1.2)	23.0 (1.2)	24.0 (1.6)	24.3 (1.7)
United States (S)	22.9 (1.3)	22.6 (1.3)	22.6 (1.2)	22.6 (1.2)	23.1 (1.5)	23.9 (1.8)	23.9 (1.5)
Qatar	22.3 (1.5)	21.8 (1.4)	22.0 (1.7)	21.9 (1.6)	21.8 (1.8)	23.0 (1.7)	23.1 (2.3)
United Kingdom	21.9 (1.4)	21.7 (1.1)	21.8 (0.9)	21.8 (0.9)	21.9 (0.9)	22.3 (1.5)	22.7 (1.6)
SOL							
Australia	18 (25)	16 (18)	17 (20)	17 (19)	16 (19)	16 (20)	16 (20)
United States (S)	14 (19)	14 (12)	16 (13)	14 (12)	12 (10)	12 (15)	10 (11)
Qatar	36 (46)	34 (56)	28 (33)	32 (43)	28 (31)	26 (36)	34 (68)
United Kingdom	30 (33)	30 (32)	32 (31)	29 (33)	30 (32)	27 (29)	27 (31)
Sleep duration							
Australia	506 (89)	498 (80)	494 (79)	495 (78)	494 (81)	520 (100)	517 (112)
United States	462 (104)	450 (88)	454 (86)	444 (85)	450 (86)	521 (115)	516 (125)
United States (S)	442 (95)	448 (101)	446 (95)	462 (104)	443 (74)	493 (112)	498 (130)
Qatar	462 (134)	462 (104)	466 (108)	459 (108)	464 (95)	506 (134)	523 (116)
United Kingdom	517 (87)	512 (71)	507 (66)	504 (65)	505 (63)	536 (83)	547 (94)
WASO							
Australia	4 (10)	7(14)	6 (14)	5 (15)	5 (17)	7 (19)	6 (15)
Wake time							
Australia	7.5 (1.4)	7.2 (1.0)	7.1 (1.0)	7.1 (0.9)	7.2 (1.7)	8.4 (1.6)	8.6 (1.6)
United States	7.0 (1.6)	6.6 (1.5)	6.6 (1.1)	6.5 (1.0)	6.6 (1.1)	8.6 (1.7)	8.9 (1.7)
United States (S)	6.9 (1.1)	6.5 (0.7)	6.4 (1.0)	6.5 (0.8)	6.6 (1.3)	8.4 (1.7)	8.5 (1.5)
Qatar	6.4 (2.2)	5.6 (2.0)	5.8 (1.2)	5.6 (1.3)	5.7 (1.2)	7.6 (1.2)	8.4 (2.2)
United Kingdom	7.2 (1.6)	6.9 (1.2)	6.9 (0.7)	6.8 (0.6)	6.9 (0.6)	7.7 (0.7)	8.4 (1.4)

School nights are Saturday to Wednesday nights for Qatar, and Sunday to Thursday nights for Australia, United Kingdom, and United States. Weekend nights are Thursday and Friday for Qatar, and Friday and Saturday nights for Australia, United Kingdom, and United States. United States (S) indicates the subsample of 55 US adolescents. SOL, sleep onset latency; WASO, wake after sleep onset.

getting up “early” to go to school. Thus, interventions are evaluated based upon improvements in school night sleep. Random effects analyses of variance were performed to determine the absolute agreement between school night sleep diary averages in week 1 (ie, the baseline week) and those in weeks 2–12, respectively.

RESULTS

Intraindividual Stability of Sleep Estimates Across 1 Week

Average sleep diary estimates of all sleep outcomes for each country are shown in Figure 1 and Table 2. These data highlight that bedtimes and wake times were consistently later weekend nights/mornings compared to weekdays. As expected, sleep duration was also longer during weekends compared to weekdays.

The number of data points available for each ICC analysis is shown in Table 3. Results of ICC analyses to determine how many nights of sleep diary entries are required for reliable estimates of sleep are displayed in Table 4. Variability was assessed for each of the four countries for school night, weekend, and weekly sleep. Overall, five nights were the minimum number of school night sleep diary entries for reliable estimates of bedtime, SOL, and sleep duration in all samples. Four nights were sufficient for wake times in the Australian and UK samples, but not the US or Qatari samples. Inadequate reliability was generally found for estimates of weekend sleep using either one or two night diary entries. Exceptions to this were two nights of weekend bedtime estimates in the Australian adolescents and two nights of either SOL estimates or wake estimates in the UK adolescents. Adequate reliability was observed for full week estimates of bedtime, SOL, sleep duration, and wake time, with

Table 3—Data Points Available According to Country for All Sleep Diary Entries.

	School night sleep					Weekend sleep		Weekly
	1 night	2 nights	3 nights	4 nights	5 nights	1 night	2 nights	7 nights
Bedtime								
Australia	339	348	346	343	339	341	341	330
United States	290	301	294	293	290	302	302	286
United States (S)	52	53	53	52	52	55	55	52
Qatar	217	233	227	219	217	239	239	215
United Kingdom	670	741	732	723	670	723	723	643
SOL								
Australia	331	343	340	337	331	302	302	290
United States (S)	47	50	47	47	47	54	54	47
Qatar	208	228	222	211	208	233	233	205
United Kingdom	575	675	647	629	575	679	679	552
Sleep duration								
Australia	334	347	344	341	334	323	323	310
United States	249	274	264	256	249	274	274	235
United States (S)	44	45	45	44	44	52	52	42
Qatar	203	228	215	207	203	219	219	187
United Kingdom	662	735	725	717	662	738	738	649
WASO								
Australia	328	335	334	330	328	329	329	312
Wake time								
Australia	341	347	346	344	341	342	342	331
United States	300	308	304	303	300	303	303	294
United States (S)	53	54	53	53	53	55	55	53
Qatar	209	230	221	212	209	233	233	203
United Kingdom	630	698	691	686	630	704	704	620

United States (S) indicates the subsample of 55 US adolescents. SOL, sleep onset latency; WASO, wake after sleep onset.

the exceptions of weekly sleep duration in the Qatari and larger US samples adolescents and weekly wake times in the Qatari and US adolescents. Information on WASO was only available from the Australian sample, which we observed to be unreliable for all time frames.

Given the number of unreliable estimates obtained from 1 week of sleep diary data, additional analyses were conducted using the subsample of US adolescents who completed 12 weeks of sleep diary entries. These analyses revealed that 13 weekend nights of bedtimes were needed for good reliability (ICC = .62) and 17 nights were needed for excellent reliability (ICC = .75). For weekend SOL, four nights were needed for good reliability (ICC = .60) and five nights for excellent reliability (ICC = .75). Weekend sleep duration needed 8 nights for good reliability (ICC = .65) and 10 nights for excellent reliability (ICC = .75). Finally, wake time on weekends required 8 nights (ICC = .63) and 15 nights (ICC = .76) for good or excellent reliability, respectively. Weekly wake time was similarly not

reliably estimated using 1 week of sleep diary entries, however, 2 full weeks were sufficient for excellent reliability (ICC = .76).

Test–Retest Reliability of Sleep Diaries

Weekly bedtime, SOL, and sleep duration on school nights were averaged for each participant for every week where at least three nights of data were available. Three nights was chosen as the minimum required because we showed adequate reliability of three nights of sleep diary entries on school nights in this sample of US adolescents for bedtime (ICC = .71), sleep duration (ICC = .72), and SOL (ICC = .69), but not for three sleep diary estimates of wake time (ICC = .39). Thus, wake time data were not included in the analyses. Test–retest reliability results and descriptive statistics for 12 weeks of sleep diary data on school nights for US adolescents are shown in Table 5. For all analyses, each week was compared to week 1. Results indicate good or excellent test–retest reliability across 12 school weeks for bedtime, SOL, and sleep duration.

Table 4—Intraclass Coefficients for School Night, Weekend, and Weekly Sleep Estimates.

	School night sleep					Weekend sleep		Weekly
	1 night	2 nights	3 nights	4 nights	5 nights	1 night	2 nights	7 nights
Bedtime								
Australia	.56	.73 ^a	.80 ^b	.83 ^b	.87 ^b	.47	.64 ^a	.80 ^b
United States	.50	.69 ^a	.74 ^a	.79 ^b	.83 ^b	.34	.51	.79 ^b
United States (S)	.53	.52	.71 ^a	.81 ^b	.85 ^b	.41	.59	.81 ^b
Qatar	.38	.77 ^a	.73 ^a	.80 ^b	.75 ^b	.27	.43	.69 ^a
United Kingdom	.47	.47	.67 ^a	.77 ^b	.82 ^b	.22	.36	.74 ^a
SOL								
Australia	.28	.33	.56	.63 ^a	.66 ^a	.29	.45	.73 ^a
United States	.41	.63 ^a	.69 ^a	.78 ^b	.78 ^b	.40	.57	.77 ^b
Qatar	.41	.61 ^a	.68 ^a	.79 ^b	.78 ^b	.20	.34	.82 ^b
United Kingdom	.52	.72 ^a	.81 ^b	.84 ^b	.85 ^b	.49	.66 ^a	.85 ^b
Sleep duration								
Australia	.34	.53	.63 ^a	.69 ^a	.72 ^a	.19	.31	.68 ^a
United States	.23	.39	.45	.56	.60 ^a	.04	.08	.43
United States (S)	.42	.68 ^a	.72 ^a	.82 ^b	.79 ^b	.20	.33	.69 ^a
Qatar	.27	.38	.52	.62 ^a	.65 ^a	.18	.30	.58
United Kingdom	.53	.65 ^a	.79 ^b	.84 ^b	.85 ^b	.30	.47	.80 ^b
WASO								
Australia	.16	.22	.29	.38	.50	.14	.24	.55
Wake time								
Australia	.29	.39	.54	.61 ^a	.67 ^a	.41	.58	.61 ^a
United States	.20	.36	.36	.49	.55	.20	.36	.44
United States (S)	.21	.35	.39	.48	.57	.22	.37	.37
Qatar	.20	.23	.36	.48	.55	.31	.47	.41
United Kingdom	.42	.56	.67 ^a	.73 ^b	.79 ^b	.42	.59	.65 ^a

United States (S) indicates the subsample of 55 US adolescents.

^aIndicates good reliability (ICC = .60–.74).

^bIndicates excellent reliability (ICC ≥ .75).

When compared to week 1, all school week estimates showed adequate stability over time (ie, ICC > .60), except for week 7 sleep duration estimates, which were just below this cutoff (ICC = .59).

DISCUSSION

Sleep diaries are a widely used tool and can provide prospective estimates of multiple sleep parameters. Diaries are relatively quick and easy to administer and have good agreement levels with wrist actigraphy¹⁸ as well as polysomnography,¹⁹ thus making them an attractive and cost-effective tool to holistically estimate sleep. We report, for the first time, comprehensive cross-cultural findings for the number of nights sleep diary entries required to reliably estimate a range of sleep measures obtained from a combined total of 1766 diary weeks across the four adolescent samples.

School Night Sleep

Results indicate that bedtime, SOL, and sleep duration can be reliably estimated with five school nights. While reliability for some sleep parameters was obtained in some adolescent groups with fewer than four school nights, collecting more data points than the absolute minimum is a conservative approach likely to ensure reliable estimates of sleep. Four nights of sleep diary entries provided reliable estimates of school night wake times in the Australian and UK adolescents, but not the Qatari and US adolescents. Reliable estimates of school night WASO were not achieved with five nights of sleep diary entries (these data were only available from the Australian adolescents). While cross-cultural differences in sleep patterns and night-to-night stability were observed, the sleep variables showing the greatest intraindividual day-to-day variation were WASO and wake times. Although variability in WASO is to be expected in this

Table 5—Test–Retest Reliability Results and Descriptive Statistics for 12 wk of Sleep Diary Data on School Nights for 55 US Adolescents.

	Week											
	1	2	3	4	5	6	7	8	9	10	11	12
Bedtime												
ICC	N/A	.88 ^a	.87 ^a	.90 ^a	.93 ^a	.89 ^a	.85 ^a	.90 ^a	.90 ^a	.88 ^a	.88 ^a	.90 ^a
N	55	55	55	54	53	51	52	50	47	49	50	50
Mean	22:47	22:58	22:58	22:50	22:56	22:52	22:53	22:34	22:34	22:47	22:49	22:52
SD	65 min	73 min	67 min	70 min	67 min	68 min	65 min	64 min	67 min	75 min	55 min	58 min
SOL												
ICC	N/A	.83 ^a	.85 ^a	.86 ^a	.82 ^a	.82 ^a	.78 ^a	.78 ^a	.76 ^a	.77 ^a	.72 ^b	.83 ^a
N	54	53	54	51	50	49	49	48	45	47	50	47
Mean	14 min	11 min	11 min	13 min	13 min	11 min	13 min	13 min	13 min	12 min	11 min	11 min
SD	10 min	10 min	11 min	11 min	11 min	9 min	10 min	12 min	12 min	12 min	10 min	10 min
SD												
ICC	N/A	.80 ^a	.71 ^b	.87 ^a	.83 ^a	.81 ^a	.59	.85 ^a	.80 ^a	.88 ^a	.84 ^a	.82 ^a
N	53	52	51	47	49	46	44	44	44	44	44	43
Mean	7 h 25 min	7 h 34 min	7 h 39 min	7 h 38 min	7 h 38 min	7 h 37 min	7 h 51 min	7 h 31 min	7 h 43 min	7 h 35 min	7 h 43 min	7 h 30 min
SD	72 min	71 min	72 min	72 min	77 min	71 min	56 min	62 min	65 min	59 min	74 min	56 min

ICC, intraclass correlation coefficient; N, number; N/A, not applicable; SD, sleep duration; SD, standard deviation; SOL, sleep onset latency; WUT, wake up time.

^aIndicates excellent reliability (ICC ≥ .75).

^bIndicates good reliability (ICC = .60–.74).

healthy, young community sample of adolescents, the day-to-day variability in wake times was more surprising. Factors that may influence this variability include changes to school start times (eg, older students may have free lessons and are not required to arrive at school until their first lesson of the day), before school sports practice, school closure days, absenteeism, and days that students take off school. Indeed, previous research provides some support for this hypothesis, with short sleep duration and sleep deficiency significantly associated with nonattendance at school.²⁰ This may help to explain why wake time reliability was worse for US and Qatari adolescents, who obtained less sleep than their UK and Australian counterparts.

Weekend Sleep

Overall, neither one nor two nights of weekend sleep diary entries (ie, when adolescents do not have school the following day) were sufficient to estimate sleep patterns of weekends among adolescents. The only exceptions to this were found for bedtimes in the Australian adolescents, and SOL in the UK adolescents. Weekend sleep in adolescents is characterized by a higher degree of variability compared to weekdays given that sleep–wake timings can be more flexible during weekends in the absence of school attendance. Typically, adolescents show characteristic delays of bedtimes and wake times across both weekend nights,⁹ coupled with increasingly longer sleep durations as they “catch up” on sleep lost across the school week and also sleep at times more aligned to their circadian phase timing.²¹ While these factors alone introduce night-to-night

variability, weekend schedules also vary depending on sport, social, and religious commitments. Resultantly, supplementary analyses of the US subsample of 55 adolescents revealed that good reliability required a minimum of 13 weekend nights of sleep diary entries for bedtimes, four weekend nights for SOL, and eight weekend nights for sleep duration and wake times.

Weekly Sleep

Despite the known differences between school night and weekend sleep in adolescents,⁹ weekly estimates of sleep provided good-to-excellent reliability when measuring bedtimes, SOL, and sleep duration. Weekly wake time estimates varied across countries, with weekly wake times showing good reliability for Australian and UK samples, for whom wake times varied less across the school week. The US and Qatari adolescents showed larger discrepancies between school night and weekend wake times and, as such, weekly wake time estimates were not reliably estimated with one week of sleep diary data, although supplementary analysis showed that 2 full weeks were sufficient for excellent weekly estimates of wake time using the subsample of US adolescents. This wake time discrepancy was shown most markedly amongst Qatari adolescents, where wake times at weekends were considerably later (~08:00 am) compared to school days (~05:40 am). Furthermore, Qatari adolescents also showed the greatest systematic variability across wake time on weekdays, waking almost 1 hour later (06:21 am) on the first weekday compared to other school mornings (~05:40 am).

Overall, the results of the present study are largely consistent with those found by Acebo and colleagues³ when examining the reliability of pediatric actigraphy. Acebo and colleagues provided wrist actigraphy to a sample of 55 adolescents, 36 of whom (aged 14–16 years) wore wrist actigraphy across 1 week during school time. The remaining adolescent analyses focused on 1 week of school holiday sleep. While, overall, the study concluded that at least five nights of actigraphic data were required for reliable sleep estimates in the pediatric populations examined (aged 1–16 years), examination of adolescent reliability across 1 school week revealed a slightly different pattern of results. Specifically, five *school* nights were adequate for reliable estimates of sleep start time, sleep minutes, and wake minutes, but not sleep period. On the other hand, reliability of five nights combining school nights *and* weekend nights was inadequate for all sleep parameters assessed, except for wake minutes. This difference highlights the importance of considering from which night's sleep estimates are drawn.

Test–Retest Reliability of Sleep Diaries Over 12 Weeks

Good-to-excellent test–retest reliability was observed for weekly estimates of school night bedtime, sleep duration, and SOL over 12 consecutive weeks in US adolescents. These ICCs can now be used to calculate reliable change indices for application in research and clinical practice. For example, many sleep interventions for adolescents span a period of 22,23 to 10 weeks,²⁴ and some include short-term follow-ups.^{22,25,26} Thus, the coefficients provided in the current study now allow researchers to assess the clinical significance of interventions with, arguably, the field's most commonly used measure. Only one ICC did not meet reliability criterion, average sleep duration in week 7 (ICC = .59). Given that reliability of sleep duration is demonstrated both before and after this week, this divergent finding is likely to reflect idiosyncrasies of the sampling method used, rather than a systematic change in sleep patterns. For example, because participants were drawn from five schools in the same city, and participants completed sleep diaries across the same time period (first semester of the year), atypical events occurring at one school (ie, school closure), or events that affect all schools at that time period (such as a public holiday or snow storm), may add increased variability to sleep at that time. This finding has important ramifications for the interpretation of sleep diary data. Even during school times, events may impact on sleep, such as sickness, days off school, and school camps and excursions. When measuring sleep over time, such events may disproportionately affect school night averages. There are ways in which these challenges may be dealt with, including collecting sleep diary data for more nights than the minimum required for reliability, and asking adolescents to indicate whether or not they attended school on each school day. Having this information may enable such events to be appropriately identified and accounted for.

Strengths and Limitations of the Current Study

The current study benefits from a large amount of sleep diary data obtained from a large sample of adolescents over four countries, capturing different cultures and both genders. We are the first to investigate the number of nights sleep diary data

that are required to reliably estimate sleep in adolescents. Our study provides a strong message to sleep researchers, epidemiologists, clinicians, and other individuals/groups interested in sleep about diaries as a reliable tool for estimating multiple sleep outcomes. On the other hand, we acknowledge several limitations to our work. First, our findings are not likely to generalize to other groups, such as teens from other countries and those either younger or older than the present group. Additional studies are required to confirm reliability in samples from other cultures, sleep disordered populations, and those with health conditions. Indeed, a priori consideration must be given to the sleep variability in the population of interest when considering how many nights sleep diaries should be completed and clinicians and researchers would profit from calculating reliability within their own samples. The sleep variables of interest also need consideration, as reliable estimates of WASO or wake time may need 2 or more weeks of sleep diaries to ensure reliability. Furthermore, our reliability estimates for multiple sleep outcomes were restricted to term time and are therefore unlikely to be representative of other periods where sleep may be different, such as term breaks and vacations. Second, identical sleep diaries were not used across all four adolescent samples. While sleep diaries tend to enquire about the same sleep outcomes, items can be worded and thus interpreted in different ways. Third, variation in adolescent sleep outcomes is influenced by the frequency of technology use,²⁷ which we did not assess in all adolescent cohorts examined in our study. Fourth, not all of the studies included in our analyses applied the same eligibility criteria; thus the combined sample of adolescents was heterogeneous where some may have been in good mental and physical health whereas there is a possibility that there may have been others who were not. Finally, reliability of WASO estimates was limited to our Australian adolescent sample, and test–retest reliability analyses were limited to the US cohort. Thus, additional research is required to extend these findings to other groups. A general caveat is needed to note the assumption that participants prospectively complete sleep diaries daily and not at the end of the week. If sleep diary entries are completed at the same time, reliability may increase if adolescents indicate the same answers. While US and Australian adolescents also called time-stamped answering machines morning and night to confirm their bedtimes and wake times each day, UK and Qatari adolescents did not. Unless electronic diaries are used and downloaded daily this continues to be a potential limitation.

CONCLUSIONS

Based on the current findings, we are confident that several sleep parameters (bedtime, SOL, and sleep duration) can be reliably estimated using sleep diaries in adolescents, provided a sufficient number of entries are completed. Researchers and clinicians who aim to capture prospective, reliable sleep information should ensure that at least five nights of entries for school nights are obtained, although complete data are always preferable. Weekend sleep outcomes are not reliably estimated in adolescents using sleep diary entries for only two nights, possibly due to large variance in weekend sleep patterns in this age group. Further reliability analysis is needed to confirm if

WASO and wake time estimates can be reliably estimated with confidence in all populations.

REFERENCES

1. Buysse DJ, Ancoli-Israel S, Edinger JD, Lichstein KL, Morin CM. Recommendations for a standard research assessment of insomnia. *Sleep*. 2006; 29(9): 1155–1173.
2. Wohlgemuth WK, Edinger JD, Fins AI, Sullivan RJ Jr. How many nights are enough? The short-term stability of sleep parameters in elderly insomniacs and normal sleepers. *Psychophysiology*. 1999; 36(2): 233–244.
3. Acebo C, Sadeh A, Seifer R, et al. Estimating sleep patterns with activity monitoring in children and adolescents: how many nights are necessary for reliable measures? *Sleep*. 1999; 22(1): 95–103.
4. Moore M, Kirchner HL, Drotar D, Johnson N, Rosen C, Redline S. Correlates of adolescent sleep time and variability in sleep time: the role of individual and health related characteristics. *Sleep Med*. 2011; 12(3): 239–245.
5. Fuligni AJ, Hardway C. Daily variation in adolescents' sleep, activities, and psychological well-being. *J Res Adolescence*. 2006; 16(3): 353–378.
6. Arora T, Broglio E, Thomas GN, Taheri S. Associations between specific technologies and adolescent sleep quantity, sleep quality, and parasomnias. *Sleep Med*. 2014; 15(2): 240–247.
7. Short MA, Gradisar M, Lack LC, Wright HR, Dohnt H. The sleep patterns and well-being of Australian adolescents. *J Adolesc*. 2013; 36(1): 103–110.
8. Wolfson AR, Carskadon MA. Sleep schedules and daytime functioning in adolescents. *Child Dev*. 1998; 69(4): 875–887.
9. Gradisar M, Gardner G, Dohnt H. Recent worldwide sleep patterns and problems during adolescence: a review and meta-analysis of age, region, and sleep. *Sleep Med*. 2011; 12(2): 110–118.
10. Short MA, Gradisar M, Lack LC, et al. A cross-cultural comparison of sleep duration between US and Australian adolescents: the effect of school start time, parent-set bedtimes, and extracurricular load. *Health Educ Behav*. 2013; 40(3): 323–330.
11. LeBourgeois MK, Giannotti F, Cortesi F, Wolfson AR, Harsh J. The relationship between reported sleep quality and sleep hygiene in Italian and American adolescents. *Pediatrics*. 2005; 115(Suppl 1): 257–265.
12. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol*. 1991; 59(1): 12–19.
13. Fisher PL, Durham RC. Recovery rates in generalized anxiety disorder following psychological therapy: an analysis of clinically significant change in the STAI-T across outcome studies since 1990. *Psychol Med*. 1999; 29(6): 1425–1434.
14. Vøllestad J, Sivertsen B, Nielsen GH. Mindfulness-based stress reduction for patients with anxiety disorders: evaluation in a randomized controlled trial. *Behav Res Ther*. 2011; 49(4): 281–288.
15. Bohus M, Haaf B, Simms T, et al. Effectiveness of inpatient dialectical behavioral therapy for borderline personality disorder: a controlled trial. *Behav Res Ther*. 2004; 42(5): 487–499.
16. Wolfson AR, Carskadon MA, Acebo C, et al. Evidence for the validity of a sleep habits survey for adolescents. *Sleep*. 2003; 26(2): 213–216.
17. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994; 6(4): 284.
18. Arora T, Broglio E, Pushpakumar D, Lodhi T, Taheri S. An investigation into the strength of the association and agreement levels between subjective and objective sleep duration in adolescents. *PLoS ONE*. 2013; 8(8): e72406.
19. Rogers AE, Caruso CC, Aldrich MS. Reliability of sleep diaries for assessment of sleep/wake patterns. *Nurs Res*. 1993; 42(6): 368–372.

20. Hysing M, Haugland S, Stormark KM, Bøe T, Sivertsen B. Sleep and school attendance in adolescence: results from a large population-based study. *Scand J Public Health*. 2015; 43(1): 2–9.
21. Crowley SJ, Acebo C, Carskadon MA. Sleep, circadian rhythms, and delayed phase in adolescence. *Sleep Med*. 2007; 8(6): 602–612.
22. Wilhelmsen-Langeland A, Saxvig IW, Pallesen S, et al. A randomized controlled trial with bright light and melatonin for the treatment of delayed sleep phase disorder: effects on subjective and objective sleepiness and cognitive function. *J Biol Rhythms*. 2013; 28(5): 306–321.
23. Tamura N, Tanaka H. Effects of a sleep education program with self-help treatment on sleeping patterns and daytime sleepiness in Japanese adolescents: a cluster randomized trial. *Chronobiol Int*. 2016; 33(8): 1073–1085.
24. Clarke G, McGlinchey EL, Hein K, et al. Cognitive-behavioral treatment of insomnia and depression in adolescents: a pilot randomized trial. *Behav Res Ther*. 2015; 69: 111–118.
25. Bonnar D, Gradisar M, Moseley L, Coughlin A-M, Cain N, Short MA. Evaluation of novel school-based interventions for adolescent sleep problems: does parental involvement and bright light improve outcomes? *Sleep Health*. 2015; 1(1): 66–74.
26. Britton WB, Bootzin RR, Cousins JC, Hasler BP, Peck T, Shapiro SL. The contribution of mindfulness practice to a multicomponent behavioral sleep intervention following substance abuse treatment in adolescents: a treatment-development study. *Subst Abus*. 2010; 31(2): 86–97.
27. Arora T, Broglio E, Thomas GN, Taheri S. Associations between specific technologies and adolescent sleep quantity, sleep quality, and parasomnias. *Sleep Med*. 2014; 15(2): 240–247.

FUNDING

This research was funded by Action Medical Research Children's Charity, the Qatar Foundation Biomedical Research Program, Australian Research Council Discovery grant DP0881261, and NIH grant MH45945.

ACKNOWLEDGMENTS

The authors wish to thank Dr Christine Acebo for her feedback on this manuscript. We thank all parents who consented their children to participate in the study, all participating adolescents, and teaching support staff at the recruited schools. We also thank Omar M. Omar, Medical Statistician at the Clinical Research Core at Weill Cornell Medicine in Qatar, for his assistance with reformatting the sleep diary data (United Kingdom and Qatar). Additionally, we thank and acknowledge our funders: Action Medical Research Children's Charity for funding MASSES (United Kingdom); Qatar Foundation, through the Biomedical Research Program award to Weill Cornell Medicine in Qatar, for funding the QUEST study (Qatar); Australian Research Council Discovery grant DP0881261; and NIH Grant MH45945. The co-first authors, MAS and TA, contributed equally to this manuscript. This research was performed at Flinders University, Weill Cornell Medicine Qatar, Brown University, and King's College. No clinical trial is indicated.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication September, 2016
Submitted in final revised form November, 2016
Accepted for publication December, 2016
Address Correspondence to: Michelle A. Short, PhD, School of Psychology, Flinders University, GPO Box 2001, Bedford Park, South Australia 5042, Australia. Telephone: +61 8 8302 1966; Fax: +61 8 8302 1902; Email: michelle.short@flinders.edu.au

DISCLOSURE STATEMENT

None declared.