

Supplementary Material

About the accuracy and problems of consumer devices in the assessment of sleep

Mohamed S. Ameen ¹, Lok Man Cheung ², Theresa Hauser ¹, Michael Hahn ¹, Manuel Schabus ^{1,3,*}

¹University of Salzburg, Department of Psychology, Laboratory for Sleep, Cognition and Consciousness Research, Hellbrunner Strasse 34, 5020 Salzburg.

²School of Psychology, University of Surrey, Surrey, United Kingdom, Stag Hill, Guildford GU2 7XH, UK.

³Center for Cognitive Neuroscience Salzburg (CCNS), Hellbrunner Strasse 34, 5020 Salzburg.

* Correspondence: manuel.schabus@sbg.ac.at; Tel: (+43) 66280445113

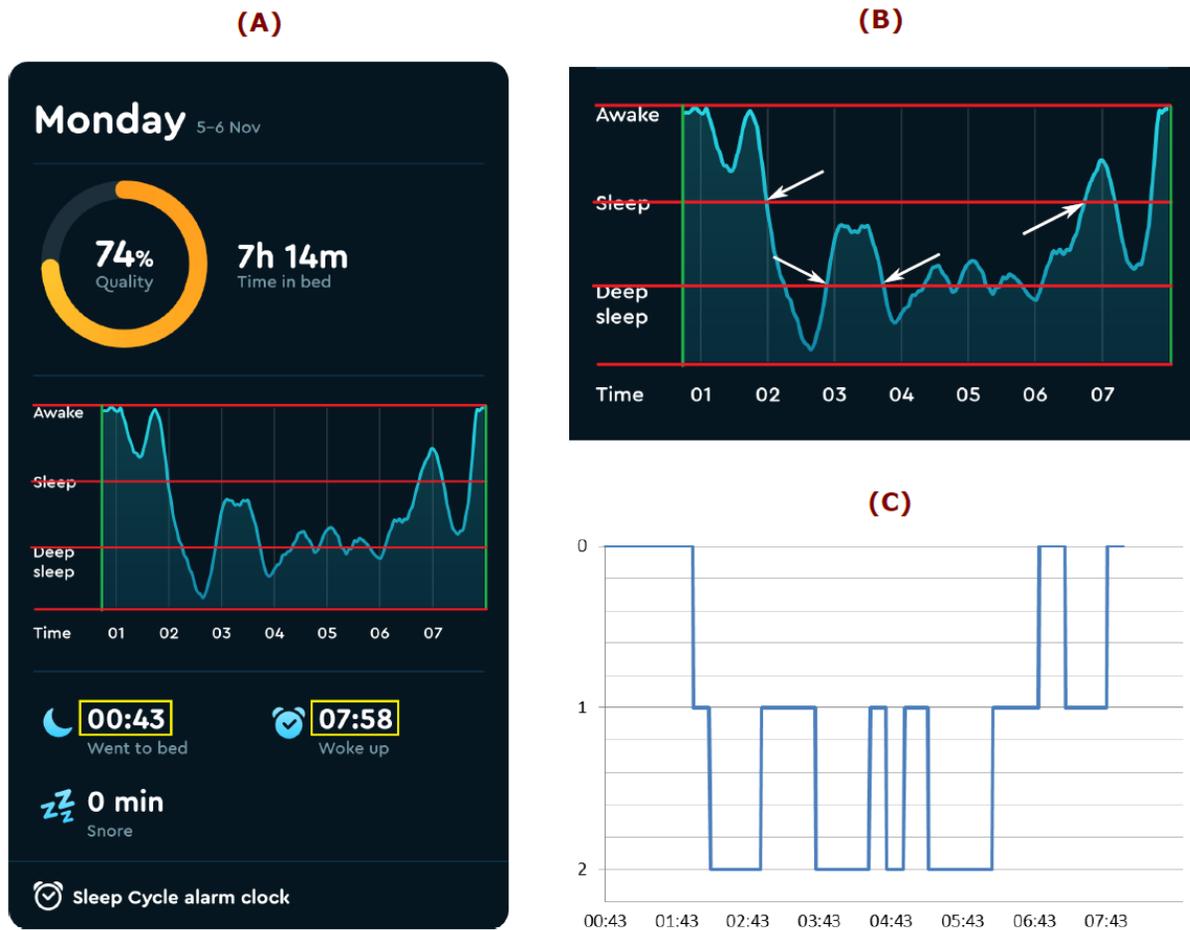


Figure S1. Sleep scoring procedure of the sleep cycle application output hypnogram. A) First we defined the boundaries of the generated hypnogram, i.e. Left, right, upper and lower boundaries, using the pixel-wise coordinates. The start and end time of the hypnogram (green lines) were defined as the time the participants went to bed and the time when they woke up respectively as reported by the application (yellow boxes). These times were then used to define the boundaries of the 30s epochs. More specifically, we calculated the overall duration of the hypnogram in minutes (time from going to bed till waking up), then calculated the total number of pixels between the left and the right boundaries of the hypnogram. Then we determine the number of pixels in 30s and these numbers were used to divide the hypnogram into 30s epochs. Furthermore, the upper and lower boundaries of the hypnogram were used to define the limits of the different sleep stages such that the distance between the upper and lower boundaries is divided by the number of stages to give the same area for each sleep stage in pixels. B) A switch from one stage to the other is detected when the tracker in the hypnogram crosses the boundary between two stages (white arrows as examples) to generate the classical 30s-epoch hypnogram depicted in

(C). This analysis was performed using image processing tools in Microsoft Paint software (Microsoft Office Professional Plus 2010) and the Visual Basic for Applications (VBA) software in Microsoft Excel (Microsoft Office Professional Plus 2010).

Time in Bed (TiB)

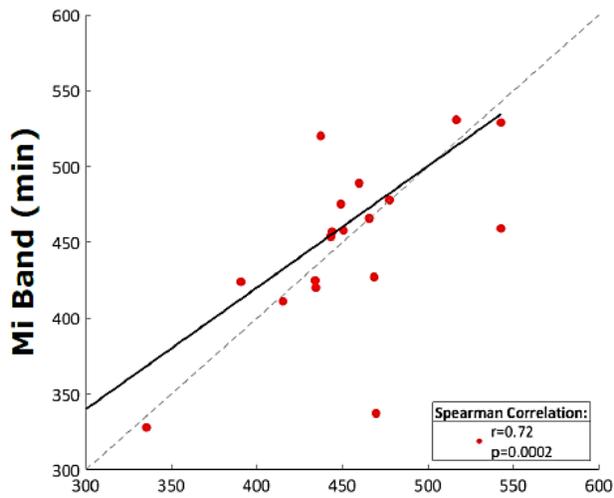


Figure S2. Scatter plot showing the positive correlations between the time in bed measurement of the PSG gold standard and those of the Mi Band, MotionWatch and sleep cycle (top to bottom). The solid black line depicts the line of correlation between the two measurements while the grey dashed line depicts the 45° line of identity.

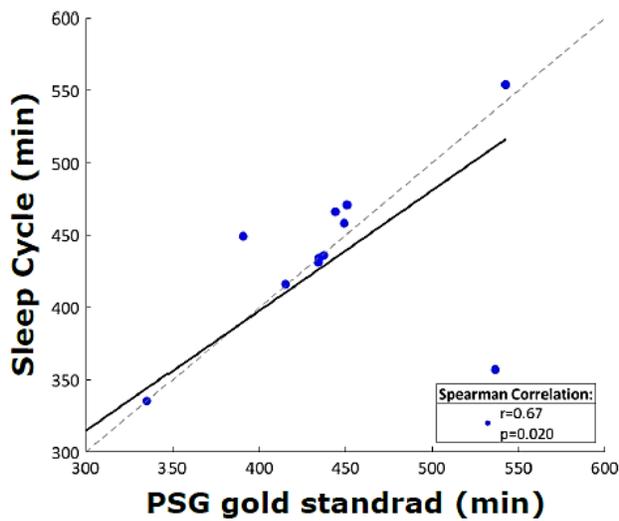
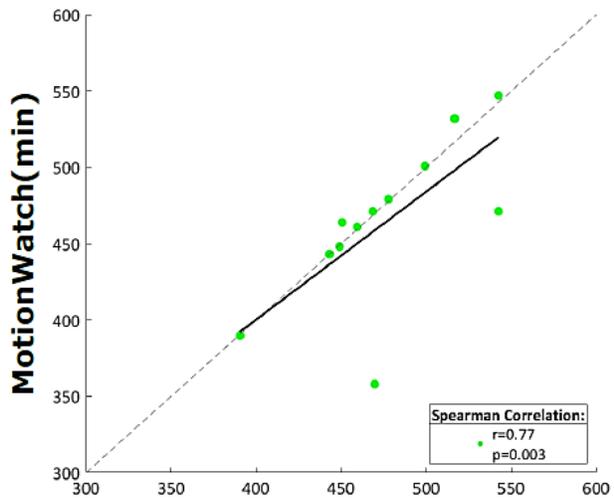


Table S1. Spearman correlation results between the PSG gold standard and the Mi Band, the MotionWatch and the Sleep Cycle application for the key sleep parameters.

Parameter	Mi Band	MotionWatch	Sleep Cycle
SOL	$r_{19} = -0.06$ $p = 0.78$	$r_{10} = -0.09$ $p = 0.76$	$r_{10} = 0.34$ $p = 0.27$
WASO	$r_{19} = 0.28$ $p = 0.22$	$r_{10} = 0.78$ $p = 0.02$	$r_{10} = 0.41$ $p = 0.17$
SE	$r_{19} = 0.43$ $p = 0.052$	$r_{10} = 0.33$ $p = 0.28$	$r_{10} = 0.32$ $p = 0.29$
TST	$r_{19} = 0.49$ $p = 0.02$	$r_{10} = 0.41$ $p = 0.17$	$r_{10} = 0.27$ $p = 0.39$
TiB	$r_{19} = 0.72$ $p = 0.0002$	$r_{10} = 0.77$ $p = 0.03$	$r_{10} = 0.67$ $p = 0.02$

SOL: Sleep onset latency, WASO: wake after sleep onset, SE: sleep efficiency and TST: total sleep time. TiB: time in bed. In bold are the significant correlations.

Table S2. Epoch-wise agreement Sleep/wake while discarding epochs scored as REM by the PSG gold standard. Note the difference in the OA and K scores as compared to Table 2 in the main article.

	PSG gold standard	
	WAKE	SLEEP
Mi Band (MB) staging		
Wake		
% Sensitivity	5.5	0.6
% PPV	62.8	37.2
Sleep		
% Sensitivity	94.5	99.4
% PPV	15.2	84.8
Sleep Cycle (SC) staging		
Wake		
% Sensitivity	55.6	30.1
% PPV	24.3	75.7
Sleep		
% Sensitivity	44.4	69.9
% PPV	9.9	90.1
MotionWatch (MW) staging		
Wake		
% Sensitivity	35.7	8.7
% PPV	43.8	56.2
Sleep		
% Sensitivity	64.3	91.3
% PPV	11.8	88.2
Device/Application	OA (%)	K
Mi Band MB	84.54	0.08
Sleep Cycle SC	67.79	0.17
MotionWatch MW	81.92	0.32

Table S3. Epoch-wise agreement comparison between (Wake/light Sleep/Deep sleep) and (Sleep/wake).

	PSG gold standard			
	WAKE	LIGHT SLEEP	DEEP SLEEP	SLEEP
MiBand (MB) staging				
Wake				
% Sensitivity	5.5	0.1	1.5	0.6
% PPV	62.8	4.7	32.6	37.2
Light sleep (Sleep)				
% Sensitivity	79.2	70.6	51.3	99.4
% PPV	18.9	57.8	23.2	84.8
Deep Sleep				
% Sensitivity	15.3	29.3	47.2	
% PPV	7.5	48.9	43.6	
Sleep Cycle (SC) staging				
Wake				
% Sensitivity	55.6	37.0	16.9	30.1
% PPV	24.3	61.1	14.7	75.7
Light sleep (Sleep)				
% Sensitivity	36.4	40.9	31.1	69.9
% PPV	14.4	61.2	24.4	91
Deep Sleep				
% Sensitivity	8.0	22.1	52.0	
% PPV	4.1	42.8	53.0	
MotionWatch (MW) staging				
Wake				
% Sensitivity	35.7	--	--	8.7
% PPV	43.8	--	--	56.2
Sleep				
% Sensitivity	64.3	--	--	91.3
% PPV	11.8	--	--	88.2

Device/Application	(W/LS/DS) OA (%) / K	(W/S) OA (%) / K
mi-band MB	53.02 / 0.17	86.54 / 0.08
Sleep Cycle SC	46.34 / 0.18	65.90 / 0.13
MotionWatch MW	--	83.42 / 0.33

We pooled Light and Deep sleep into one category labelled ‘Sleep’ and measured the agreement when the devices are only required to assign each epoch to one of two categories (Wake/Sleep) instead of 3 categories (Wake/Light sleep/Deep sleep).