Supplemental Methods

Sample Differences

The Midwestern and the Southwestern samples differed on gender (χ^2 (1, N = 1566) = 35.99, p < .001), with the Southwestern sample disproportionately female. They also differed on race/ethnicity (χ^2 (4, N = 1585) = 295.90, p < .001), with the Midwestern sample more heavily White and less heavily Asian/Asian-American, Hispanic/Latinx, and Multiracial/Other. A t-test also suggested a significant difference in age, t (1547) = 11.40, p < .001, with the Midwestern sample younger.

Close-Ended Reports of Use

The open-ended and close-ended questions on SNS use were counterbalanced and separated with questions relevant to alternative aims of the study to diminish the likelihood of prior question recall. The close-ended questions were on a 16-point scale with responses in 15-minute increments, and asked, "Approximately how many minutes do you spend per day using the [platform] app on your mobile phone, overall?" (1 = 0-15 minutes, 16 = 225-240 minutes). We opted for 15-minute intervals to strike a balance between a) being as precise in our close-ended assessments as possible (larger intervals such as hour-long increments would be less precise) while b) not overloading participants with a wealth of possible options. Further, by focusing on precision, we aimed to more closely parallel the open-ended nature of both the open-ended self-reports and device-reports of use.

Alternative Methods to Accessing Platforms

For Facebook, Instagram, and Twitter, participants were asked, "Do you ever access [platform] through a mode other than the [platform] app? (e.g., a computer or your phone's web browser)?". This item was not asked for Snapchat, as this platform can only be accessed through phone applications.

Well-being

Participants completed three assessments of well-being: self-esteem, life satisfaction, and depressive symptoms. Participants completed the 10-item Rosenberg Self-Esteem Scale (Rosenberg, 1965), answering items on a 4-point Likert-type scale (1 = Strongly Agree, 4 = Strongly Disagree; e.g., "On the whole, I am satisfied with myself", α = .88; recoded so that higher scores indicate higher self-esteem). One item was removed due to a typo; the inclusion of this item did not change the substantive pattern of results. Participants also completed the 5-item Satisfaction with Life Scale (Diener et al., 1985), answering items on a 7-point Likert-type scale (1 = Strongly Disagree, 7 = Strongly Agree; e.g., "I am satisfied with life", α=.88). Participants completed the 9-item Patient Health Questionnaire (Kroenke & Spitzer, 2002), answering how frequently they experienced depressive symptoms over the last two weeks using a 4-point Likert-type scale (e.g., 1 = Not at all, 4 = Nearly every day; "Feeling down, depressed, or hopeless", α = .90). A confirmatory factor analysis was run to establish that all items loaded on their respective scales. Standardized factor loadings were satisfactory for the models for self-esteem (loadings > .44), life satisfaction (loadings > .65) and depressive symptoms (loadings > .60).

Supplemental Results

Descriptive statistics are in Table S1. Correlations between device-reports, open-ended self-reports, and close-ended self-reports are reported in Table S2. Figure S1 demonstrates the pattern of overestimating, underestimating, and accurate reporting (defined as within 10% of the device-reported value; Ernala et al., 2020), with most participants overestimating their use for each platform. Table S3 depicts correlations between device-reported SNS use and study moderators. Because self-reports and Screen Time reports of use only assess SNS app use, it is important to note the frequencies with which participants access SNSs outside of their SNS apps, as this could influence their accuracy of reporting. Accordingly, we found that 47% of Facebook

users reported accessing Facebook via a mode other than (or in addition to) the Facebook app on their iPhone; 27% of Twitter users reported this for Twitter, and 15% of Instagram users reported this for Instagram. Of the 47% who reported using Facebook via another mode, 80% reported using this mode less than 15 minutes per day. Of the 27% who reported using Twitter via another mode, 77% reported using this mode less than 15 minutes per day. Of the 15% who reported using Instagram via another mode, 81% reported using this mode less than 15 minutes per day. In all, most participants solely accessed these platforms via their iPhone apps; for the minority who reported using another mode, access was generally low.

Full sensitivity analyses are reported for the Truth and Bias models in Tables S4-S11. Given the substantial skewness of the data and the presence of outliers, we decided to run additional quantile regression analyses (we thank an anonymous reviewer for this suggestion) for the central Truth and Bias model analyses (Tables S12-S15). The quantile regression analyses permitted us to explore how the results varied as a function of the quantile of the outcome variable. In particular, we inspected the results for the .25, .50, and .75 quantiles of the outcomes. Tables S12-15 show that tracking accuracy for each platform was positive and prominent across all the quantiles; in contrast, participants underestimated their use at the .25 quantile (though only significantly so for Facebook use) and over-estimated their use at the .50 and .75 quantiles. Besides this, little evidence emerged for any individual difference variable consistently moderating tracking accuracy or directional bias across the quantiles.

Associations between SNS use and well-being are in Table S16. Out of 24 possible associations with self-reports, nine were significant (three with the more stringent p-value of .01). These associations were small using Cohen's (1992) standards, and all were in the direction in which greater use was associated with poorer well-being. Moreover, effect sizes were

comparable with previous meta-analytic evidence suggesting trivial associations between SNS use and well-being (e.g., Huang, 2017). The most consistent association was with Twitter use. In contrast, out of the 12 possible associations with device-reports, only two were significant (neither with the more stringent p-value). These associations were also small, but in the opposite direction, in which greater use (for Instagram and Snapchat only) was associated with greater well-being.

In replication of Sewall et al. (2020), additional Truth and Bias models examined if depressive symptoms and life satisfaction moderated tracking accuracy and directional bias for each of the four platforms (Tables S17-S20). For Instagram, there was a consistent moderating effect of depressive symptoms on directional bias, such that participants displayed significantly less over-estimation of Instagram use when they were lower in depressive symptoms (i.e., 1 SD below the mean; $b = 11.97$, $SE = 5.82$, $p = .040$) compared to when they were higher in depressive symptoms (i.e., 1 SD above the mean; $b = 20.77$, $SE = 2.98$, $p < .001$). Likewise, for Twitter, there was a consistent moderating effect of depressive symptoms on directional bias, such that participants displayed significantly less over-estimation of Twitter use when they were lower in depressive symptoms (i.e., 1 SD below the mean; $b = 6.26$, $SE = 6.11$, $p = .306$) compared to when they were higher in depressive symptoms (i.e., 1 SD above the mean; $b = 16.89$, $SE = 3.20$, $p < .001$). It is notable that findings were only robust for Instagram and Twitter, as these two platforms demonstrated stronger evidence in the relation between self-reported use and poorer well-being (whereas associations for Facebook and Snapchat were not significant using the more stringent p-value). Thus, findings could suggest that previous associations between SNS use and poorer well-being could be at least partially driven by inaccurate reporting.

Supplemental References

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, 49, 71-75.

Ernala, S. K., Burke, M., Leavitt, A., & Ellison, N. B. (2020). How well do people report time spent on Facebook? An evaluation of established survey questions with recommendations. In *Proceedings of the 2020 ACM CHI Conference on Human Factors in Computing Systems,.* 1-14. doi: 10.1145/3313831.3376435

Huang, C. (2017). Time spent on social network sites and psychological well-being: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking, 20*. doi: 10.1089/cyber.2016.0758

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals, 32*, 1-7. doi: 10.3928/0048-5713-20020901-06

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Sewall, C. J. R., Bear, T. M., Merranko, J., & Rosen, D. (2020). How psychological well-being and usage amount predict inaccuracies in retrospective estimates of digital technology use. *Mobile Media & Communication,* 1-20. Advance Online Publication. doi: 10.1177/2050157920902830

Table S1

Descriptive Statistics for Study Variables

	M (SD)	Scale Range
Facebook (*n* = 1024)		
Open-Ended	26.81 (39.82)	0 – 400
Close-Ended	2.52 (2.64)	1 – 16
Device-reported	12.81 (33.47)	0 – 441
Instagram (*n* = 1403)		
Open-Ended	59.51 (55.14)	0 – 500
Close-Ended	4.67 (3.27)	1 – 16
Device-reported	30.75 (36.47)	0 – 457
Twitter (*n* = 924)		
Open-Ended	47.70 (50.82)	0 – 500
Close-Ended	4.04 (3.40)	1 – 16
Device-reported	21.29 (30.36)	0 – 307
Snapchat (*n* = 1402)		
Open-Ended	58.64 (56.80)	0 – 500
Close-Ended	4.73 (3.71)	1 – 16
Device-reported	29.53 (35.29)	0 – 286
Life Satisfaction	4.48 (1.34)	1 – 7
Self-Esteem	2.85 (0.56)	1 – 4
Depressive Symptoms	1.85 (0.67)	1 – 4
Screen Time Checking	0.12 (0.80)	0 – 26
Conscientiousness	3.46 (0.63)	1 – 5

Note. *N*'s represent number of users for each platform. All open-ended and device-reported measurements are daily and reported in minutes. For ease of interpretation for the close-ended reports, please note that the 16-point scales are in 15-minute increments (1 = 0-15 minutes per day, 2 = 15-30 minutes per day, 3 = 30-45 minutes per day, 4 = 45-60 minutes per day, 5 = 60-75 minutes per day, etc.).

Table S2

Partial Correlations between Different Types of Reports of SNS Use

	Device-Reported	Close-Ended
Facebook		
Open-Ended	.52**	.51**
Close-Ended	.35**	
Instagram		
Open-Ended	.42**	.65**
Close-Ended	.30**	
Twitter		
Open-Ended	.59**	
Close-Ended	.47**	.66**
Snapchat		
Open-Ended	.44**	
Close-Ended	.32**	.70**

Note. Recruitment site is controlled for in analyses. Open-ended and device-reported values were log-transformed because of skewness. **$p < .001$

Figure S1

Patterns of Underestimating, Overestimating, and Accurate Reporting
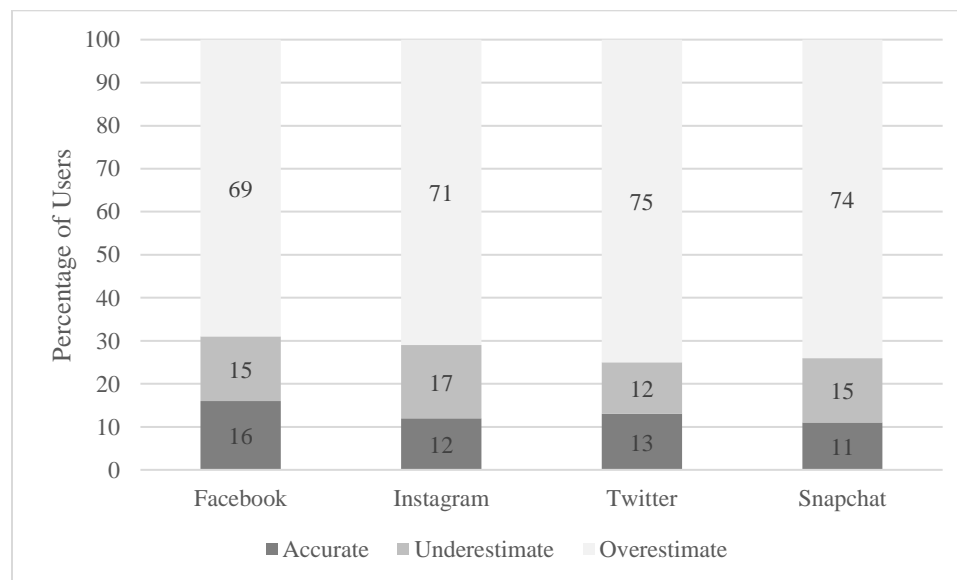
Note. Values were created by computing difference scores in which device-reports were subtracted from open-ended self-reports for each platform. Negative difference scores represent underestimating and positive difference scores represent overestimating. Accurate scores were defined as the self-report being within 10% of the device-report. Frequencies of each type of difference score (i.e., negative/underestimate, positive/overestimate, accurate) were subsequently divided by the total number of users with self-reports and device-reports for each platform.

Table S3

Partial Correlations between Device-Reported SNS use and Study Moderators

	1	2	3	4	5	6	7
1. Facebook	--						
2. Instagram	.39**	--					
3. Twitter	.36**	.56**	--				
4. Snapchat	.38**	.72**	.58**	--			
5. Age	.07*	-.10**	-.08*	-.17**	--		
6. Cons	.02	.06*	-.02	.03	.09**	--	
7. STA Check	.01	.01	.01	.00	-.01	-.01	--

Note. Recruitment site is controlled. Device-reported SNS use is log-transformed. Cons = Conscientiousness. STA Check = Screen Time App Checking.
**p < .01, *p < .05

Robustness Checks for Primary Truth and Bias Model Analyses

Table S4

Truth and Bias Model Analyses for Facebook Use Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Average Bias and Accuracy										
Average Directional Bias	13.94**	1.41	13.94**	1.43	18.20**	1.77	7.97**	0.67	11.02**	0.88
Average Tracking Accuracy	0.53**	0.08	0.53**	0.15	0.48**	0.09	0.58**	0.04	0.81**	0.07
Moderators of Directional Bias										
Site of Data Collection	-1.05	1.30	-1.05	1.36	-1.26	1.69	0.05	0.62	-0.63	0.78
Gender	-0.90	1.40	-0.90	1.36	-1.98	1.78	0.43	0.67	-0.29	0.86
Age	0.37	0.36	0.37	0.37	0.53	0.53	0.52**	0.17	0.46	0.26
Conscientiousness	-1.38	1.96	-1.38	2.12	-4.43	2.58	0.45	0.93	1.20	1.18
Screen Time App Checking	2.17	4.09	2.17	5.02	0.29	4.47	4.08*	1.94	2.04	3.13
Moderators of Tracking Accuracy										
Site of Data Collection	0.07	0.04	0.07	0.10	0.06	0.05	0.12**	0.02	0.01	0.05
Gender	0.27**	0.07	0.27	0.14	0.26**	0.08	0.21**	0.03	0.16*	0.06
Age	0.01	0.02	0.01	0.05	0.01	0.02	0.02**	0.01	-0.02	0.02
Conscientiousness	0.17*	0.07	0.17	0.13	0.18*	0.08	0.23**	0.04	0.36**	0.08
Screen Time App Checking	0.39	0.38	0.39	0.64	0.41	0.45	0.46*	0.18	0.32	0.32

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. Gender was effect-coded, such that -1 = women and 1 = men. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *p < .05. **p < .01.

Table S5

Truth and Bias Model Analyses for Instagram Use Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Average Bias and Accuracy										
Average Directional Bias	27.50**	1.52	27.50**	1.55	26.65**	1.61	19.69**	1.01	21.68**	1.15
Average Tracking Accuracy	0.69**	0.05	0.69**	0.06	0.73**	0.05	0.64**	0.03	0.61**	0.04
Moderators of Directional Bias										
Site of Data Collection	0.68	1.45	0.68	1.42	-1.17	1.56	0.25	0.96	0.27	1.07
Gender	-3.53*	1.52	-3.53*	1.53	-3.44*	1.61	-3.19**	1.01	-4.75**	1.15
Age	0.06	0.64	0.06	0.84	-0.10	0.67	-0.56	0.43	-0.33	0.54
Conscientiousness	-6.26**	2.17	-6.26**	2.29	-5.96*	2.34	-4.37**	1.44	-6.21**	1.60
Screen Time App Checking	-1.30	2.73	-1.30	3.80	-1.18	2.69	-0.04	1.81	-4.42	2.85
Moderators of Tracking Accuracy										
Site of Data Collection	0.03	0.04	0.03	0.05	0.08	0.05	0.03	0.03	-0.00	0.04
Gender	0.15**	0.05	0.15**	0.06	0.15**	0.05	0.07*	0.03	0.08*	0.04
Age	0.04	0.02	0.04	0.03	0.04	0.02	0.01	0.01	0.03	0.02
Conscientiousness	-0.05	0.07	-0.05	0.07	-0.06	0.07	-0.06	0.04	-0.06	0.05
Screen Time App Checking	-0.01	0.11	-0.01	0.12	-0.03	0.11	0.03	0.07	0.06	0.09

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. Gender was effect-coded, such that -1 = women and 1 = men. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *$p < .05$. **$p < .01$.

Table S6

Truth and Bias Model Analyses for Twitter Use Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Average Bias and Accuracy										
Average Directional Bias	24.92**	1.63	24.92**	1.70	25.47**	1.93	18.49**	1.03	20.25**	1.13
Average Tracking Accuracy	0.83**	0.06	0.83**	0.07	0.81**	0.07	0.82**	0.04	0.85**	0.05
Moderators of Directional Bias										
Site of Data Collection	-1.44	1.60	-1.44	1.64	-1.88	1.89	0.31	1.01	0.76	1.11
Gender	-3.70*	1.62	-3.70*	1.62	-3.92*	1.90	-3.74**	1.02	-4.46**	1.14
Age	-0.65	0.65	-0.65	0.60	-1.27	1.03	-0.44	0.41	-0.86	0.54
Conscientiousness	-4.74	2.42	-4.74	2.47	-5.11	2.88	-4.79**	1.52	-4.36**	1.68
Screen Time App Checking	0.08	2.41	0.08	2.77	-0.32	2.50	-0.80	1.52	-3.12	3.47
Moderators of Tracking Accuracy										
Site of Data Collection	0.00	0.06	0.00	0.08	-0.00	0.06	0.07*	0.04	0.11*	0.05
Gender	-0.03	0.06	-0.03	0.08	-0.03	0.06	-0.10**	0.04	-0.05	0.06
Age	0.01	0.02	0.01	0.03	0.03	0.03	0.02	0.01	-0.00	0.02
Conscientiousness	0.05	0.07	0.05	0.09	0.05	0.08	0.02	0.05	0.03	0.07
Screen Time App Checking	0.08	0.17	0.08	0.19	0.02	0.18	0.00	0.10	0.15	0.15

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. Gender was effect-coded, such that -1 = women and 1 = men. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *p < .05. **p < .01.

Table S7

Truth and Bias Model Analyses for Snapchat Use Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Average Bias and Accuracy										
Average Directional Bias	27.11**	1.61	27.11**	1.82	24.80**	1.61	20.72**	1.05	22.99**	1.22
Average Tracking Accuracy	0.68**	0.05	0.68**	0.09	0.74**	0.05	0.70**	0.03	0.73**	0.05
Moderators of Directional Bias										
Site of Data Collection	1.76	1.51	1.76	1.60	-0.35	1.53	0.85	0.99	1.26	1.11
Gender	-1.95	1.53	-1.95	1.56	-1.06	1.56	-1.03	1.00	-2.13	1.12
Age	-1.96*	0.83	-1.96	1.31	-2.04*	0.81	-1.24*	0.54	-1.66*	0.68
Conscientiousness	-2.30	2.23	-2.30	2.59	-0.63	2.27	-0.83	1.46	-1.45	1.63
Screen Time App Checking	4.25	4.37	4.25	6.44	3.53	4.17	2.71	2.86	-1.84	4.23
Moderators of Tracking Accuracy										
Site of Data Collection	0.01	0.04	0.01	0.08	0.06	0.04	0.01	0.03	-0.04	0.04
Gender	0.05	0.04	0.05	0.08	0.03	0.04	0.01	0.03	0.04	0.04
Age	-0.01	0.03	-0.01	0.06	0.00	0.03	0.01	0.02	0.01	0.03
Conscientiousness	0.01	0.07	0.01	0.16	-0.03	0.07	0.03	0.04	0.01	0.05
Screen Time App Checking	0.23	0.21	0.23	0.31	0.14	0.20	0.12	0.14	0.29	0.21

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. Gender was effect-coded, such that -1 = women and 1 = men. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *$p < .05$. **$p < .01$.

Robustness Checks for Primary Truth and Bias Model Analyses with Quadratic Effects for Tracking Accuracy

Table S8

Truth and Bias Model Analyses for the Quadratic Effects of Facebook Use Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Covariates										
Site of Data Collection	-1.941	1.147	-1.941	1.168	-2.510	1.458	-0.994	0.511	-1.261*	0.577
Bias and Accuracy Terms										
Directional Bias	16.545**	1.177	16.545**	1.292	19.170**	1.472	9.868**	0.524	10.976**	0.613
Linear - Tracking Accuracy	0.876**	0.070	0.876**	0.103	0.809**	0.081	0.867**	0.031	0.846**	0.044
Squared – Tracking Accuracy	-0.002**	0.000	-0.002**	0.000	-0.002**	0.000	-0.002**	0.000	-0.002**	0.000

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *$p < .05$. **$p < .01$.

Table S9

Truth and Bias Model Analyses for the Quadratic Effects of Instagram Use Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Covariates										
Site of Data Collection	-0.346	1.362	-0.346	1.376	-1.788	1.464	0.071	0.894	-0.020	0.932
Bias and Accuracy Terms										
Directional Bias	29.765**	1.402	29.765**	1.468	28.522**	1.487	21.690**	0.920	23.308**	1.088
Linear - Tracking Accuracy	0.706**	0.050	0.706**	0.058	0.792**	0.056	0.667**	0.033	0.649**	0.041
Squared – Tracking Accuracy	-0.001*	0.000	-0.001*	0.000	-0.001**	0.000	-0.000*	0.000	-0.001	0.001

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *$p < .05$. **$p < .01$.

Table S10

Truth and Bias Model Analyses for the Quadratic Effects of Twitter Use Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Covariates										
Site of Data Collection	-1.937	1.469	-1.937	1.470	-2.213	1.700	-0.624	0.951	-0.203	0.925
Bias and Accuracy Terms										
Directional Bias	27.583**	1.524	27.583**	1.716	28.271**	1.718	20.319**	0.987	22.330**	1.115
Linear - Tracking Accuracy	0.950**	0.070	0.950**	0.083	0.925**	0.082	0.879**	0.046	0.964**	0.055
Squared – Tracking Accuracy	-0.001*	0.001	-0.001	0.001	-0.001	0.001	-0.001	0.000	-0.002*	0.001

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *$p < .05$. **$p < .01$.

Table S11

Truth and Bias Model Analyses for the Quadratic Effects of Snapchat Use Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Covariates										
Site of Data Collection	2.179	1.382	2.179	1.397	-0.133	1.391	1.474	0.902	2.140*	0.921
Bias and Accuracy Terms										
Directional Bias	31.212**	1.507	31.212**	1.564	28.968**	1.448	22.804**	0.984	22.921**	1.051
Linear - Tracking Accuracy	0.858**	0.058	0.858**	0.074	1.063**	0.061	0.798**	0.038	0.723**	0.041
Squared – Tracking Accuracy	-0.002**	0.000	-0.002**	0.001	-0.003**	0.000	-0.001**	0.000	-0.001	0.001

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *$p < .05$. **$p < .01$.

Quantile Regression Analyses

Table S12

Truth and Bias Model Analyses for Facebook Use: Quantile Regression Analyses at 0.25, 0.50, and 0.75

	Quantile					
	0.25		0.50		0.75	
Predictors	b	SE	b	SE	b	SE
Average Bias and Accuracy						
Average Directional Bias	-4.33**	0.64	4.91**	0.93	20.21**	2.14
Average Tracking Accuracy	0.46**	0.06	0.68**	0.09	0.90**	0.21
Moderators of Directional Bias						
Site of Data Collection	0.23	0.70	0.39	0.95	-1.73	1.93
Gender	2.19**	0.57	0.93	0.82	-2.49	1.85
Age	0.39	0.42	0.45*	0.17	0.94	0.51
Conscientiousness	1.34	0.90	1.02	1.40	-3.67	2.56
Screen Time App Checking	1.07	1.91	6.35	5.50	6.24	7.74
Moderators of Tracking Accuracy						
Site of Data Collection	0.02	0.08	0.16	0.10	0.03	0.18
Gender	0.32**	0.06	0.22**	0.08	-0.04	0.17
Age	0.03	0.04	0.04*	0.01	-0.04	0.05
Conscientiousness	0.13	0.09	0.25	0.14	0.33	0.24
Screen Time App Checking	0.14	0.21	0.73	0.45	0.60	0.99

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. Gender was effect-coded, such that -1 = women and 1 = men. All continuous predictors were grand-mean centered. $*p < .05$. $**p < .01$.

Table S13

Truth and Bias Model Analyses for Instagram Use: Quantile Regression Analyses at 0.25, 0.50, and 0.75

	Quantile					
	0.25		0.50		0.75	
Predictors	b	SE	b	SE	b	SE
Average Bias and Accuracy						
Average Directional Bias	-1.24	0.93	14.25**	1.33	39.92**	2.58
Average Tracking Accuracy	0.60**	0.04	0.62**	0.05	0.77**	0.11
Moderators of Directional Bias						
Site of Data Collection	-0.33	0.86	-0.89	1.12	0.22	2.05
Gender	-0.69	0.92	-3.09*	1.21	-5.69*	2.43
Age	-0.73**	0.22	-0.96	0.63	-0.13	1.16
Conscientiousness	-1.72	1.21	-3.69*	1.61	-7.39*	2.89
Screen Time App Checking	2.88	3.32	0.61	4.86	-2.17	9.00
Moderators of Tracking Accuracy						
Site of Data Collection	0.02	0.03	0.02	0.03	0.02	0.08
Gender	0.01	0.04	0.07	0.04	0.10	0.10
Age	-0.02	0.01	-0.01	0.02	0.03	0.05
Conscientiousness	-0.12*	0.05	-0.10*	0.04	0.09	0.11
Screen Time App Checking	0.10	0.11	0.02	0.07	-0.07	0.34

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. Gender was effect-coded, such that -1 = women and 1 = men. All continuous predictors were grand-mean centered. *$p < .05$. **$p < .01$.

Table S14

Truth and Bias Model Analyses for Twitter Use: Quantile Regression Analyses at 0.25, 0.50, and 0.75

| | Quantile | | | | | |
| | 0.25 | | 0.50 | | 0.75 | |
Predictors	b	SE	b	SE	b	SE
Average Bias and Accuracy						
Average Directional Bias	-0.31	0.78	12.61**	1.37	36.69**	2.24
Average Tracking Accuracy	0.75**	0.04	0.86**	0.06	0.97**	0.10
Moderators of Directional Bias						
Site of Data Collection	-0.11	0.71	1.23	1.39	1.14	2.13
Gender	-2.15**	0.66	-3.25**	1.25	-7.75**	2.01
Age	-0.07	0.50	0.13	0.91	-0.84	1.02
Conscientiousness	-0.66	0.98	-5.40**	2.01	-7.78*	3.24
Screen Time App Checking	1.74	4.82	0.14	2.86	-0.95	9.04
Moderators of Tracking Accuracy						
Site of Data Collection	0.01	0.04	0.09	0.07	0.08	0.10
Gender	-0.05	0.04	-0.09	0.06	-0.04	0.10
Age	0.00	0.02	0.02	0.04	0.01	0.05
Conscientiousness	0.05	0.05	0.01	0.09	0.11	0.15
Screen Time App Checking	0.15	0.14	0.06	0.21	0.04	0.43

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. Gender was effect-coded, such that -1 = women and 1 = men. All continuous predictors were grand-mean centered. *$p < .05$. **$p < .01$.

Table S15

Truth and Bias Model Analyses for Snapchat Use: Quantile Regression Analyses at 0.25, 0.50, and 0.75

	Quantile					
	0.25		0.50		0.75	
Predictors	b	SE	b	SE	b	SE
Average Bias and Accuracy						
Average Directional Bias	-1.70	0.91	13.69**	1.32	43.29**	2.03
Average Tracking Accuracy	0.63**	0.04	0.75**	0.05	0.85**	0.08
Moderators of Directional Bias						
Site of Data Collection	0.56	0.97	1.01	1.28	2.28	2.03
Gender	0.12	0.81	-1.67	1.17	-2.49	1.86
Age	-0.06	0.50	-1.24	0.70	-2.31*	1.03
Conscientiousness	1.68	1.23	-0.76	1.77	-1.97	2.99
Screen Time App Checking	5.35	5.45	2.69	4.53	1.44	9.50
Moderators of Tracking Accuracy						
Site of Data Collection	0.03	0.04	-0.02	0.05	-0.07	0.08
Gender	0.02	0.04	-0.01	0.04	0.05	0.07
Age	0.01	0.02	-0.02	0.03	-0.01	0.04
Conscientiousness	0.11*	0.06	0.02	0.06	-0.07	0.12
Screen Time App Checking	0.24	0.21	0.07	0.22	0.06	0.32

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. Gender was effect-coded, such that -1 = women and 1 = men. All continuous predictors were grand-mean centered. *$p < .05$. **$p < .01$.

Table S16

Partial Correlations Between Reports of SNS Use and Well-being

	Life Satisfaction	Self-Esteem	Depressive Symptoms
Facebook			
Open	.02 (.02)	-.02 (-.02)	.04 (.01)
Close	-.04 (-.02)	-.07* (-.05*)	.07* (.04)
Device	.02 (.03)	.00 (-.01)	-.01 (-.01)
Instagram			
Open	.02 (.03)	.00 (-.02)	.04 (.04)
Close	-.01 (.00)	-.03 (-.04)	.09** (.08**)
Device	.07* (.07**)	.04 (.02)	-.04 (-.02)
Twitter			
Open	-.04 (-.02)	-.07* (-.08**)	.10** (.11**)
Close	-.08* (-.05*)	-.08* (-.09**)	.12** (.12**)
Device	.01 (.00)	.00 (-.05)	.02 (.07*)
Snapchat			
Open	.04 (.05)	.01 (-.01)	.01 (.02)
Close	-.01 (.01)	-.02 (-.02)	.06* (.06*)
Device	.06* (.07**)	.02 (.01)	-.02 (-.01)

Note. Recruitment suit was controlled. Due to skewness, open-ended and device-reported values were log transformed. Values in parentheses represent associations including participants without an account on a given platform, in which use was scored as '0' for all types of reports.
**p<.01, *p<.05

Robustness Checks for Supplemental Truth and Bias Model Analyses with Depression and Life Satisfaction as Moderators of Tracking Accuracy and Directional Bias

Table S17

Truth and Bias Model Analyses for Facebook Use with Depression and Life Satisfaction as Moderators of Tracking Accuracy and Directional Bias Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Average Bias and Accuracy										
Average Directional Bias	13.72**	1.20	13.72**	1.27	18.71**	1.54	7.71**	0.57	10.41**	0.70
Average Tracking Accuracy	0.27**	0.04	0.27*	0.13	0.22**	0.04	0.46**	0.02	0.66**	0.05
Moderators of Directional Bias										
Site of Data Collection	-1.97	1.23	-1.97	1.36	-2.34	1.56	-0.85	0.58	-1.21	0.71
Depression	3.87	2.08	3.87	2.67	4.72	2.68	0.76	0.99	1.47	1.23
Life Satisfaction	1.27	1.03	1.27	1.37	0.52	1.33	0.55	0.49	0.17	0.61
Moderators of Tracking Accuracy										
Site of Data Collection	0.04	0.04	0.04	0.12	0.04	0.04	0.05**	0.02	0.02	0.05
Depression	-0.05	0.07	-0.05	0.25	-0.06	0.08	-0.08*	0.03	0.16	0.09
Life Satisfaction	0.05	0.04	0.05	0.09	0.05	0.04	0.06**	0.02	0.03	0.04

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *$p < .05$. **$p < .01$.

Table S18

Truth and Bias Model Analyses for Instagram Use with Depression and Life Satisfaction as Moderators of Tracking Accuracy and Directional Bias Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Average Bias and Accuracy										
Average Directional Bias	28.58**	1.34	28.58**	1.34	27.54**	1.45	21.12**	0.89	23.04**	0.95
Average Tracking Accuracy	0.62**	0.04	0.62**	0.05	0.65**	0.04	0.61**	0.02	0.64**	0.03
Moderators of Directional Bias										
Site of Data Collection	-0.28	1.36	-0.28	1.38	-2.13	1.47	0.07	0.90	-0.28	0.96
Depression	6.60**	2.24	6.60**	2.51	7.28**	2.41	4.90**	1.49	6.01**	1.63
Life Satisfaction	1.63	1.15	1.63	1.27	2.30	1.25	0.58	0.76	1.70*	0.83
Moderators of Tracking Accuracy										
Site of Data Collection	0.02	0.04	0.02	0.05	0.07	0.04	0.03	0.02	0.00	0.03
Depression	0.06	0.07	0.06	0.09	0.04	0.08	0.09	0.05	0.15**	0.06
Life Satisfaction	0.03	0.03	0.03	0.04	0.03	0.04	0.02	0.02	0.06	0.03

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *$p < .05$. **$p < .01$.

Table S19

Truth and Bias Model Analyses for Twitter Use with Depression and Life Satisfaction as Moderators of Tracking Accuracy and Directional Bias Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Average Bias and Accuracy										
Average Directional Bias	26.32**	1.48	26.32**	1.52	27.21**	1.73	19.89**	0.94	20.91**	0.96
Average Tracking Accuracy	0.85**	0.06	0.85**	0.07	0.82**	0.06	0.84**	0.04	0.87**	0.04
Moderators of Directional Bias										
Site of Data Collection	-2.41	1.51	-2.41	1.54	-2.99	1.74	-0.85	0.96	-0.58	0.98
Depression	7.97**	2.32	7.97**	2.21	8.66**	2.69	7.05**	1.47	8.17**	1.53
Life Satisfaction	1.59	1.25	1.59	1.14	2.51	1.49	1.14	0.79	2.38**	0.82
Moderators of Tracking Accuracy										
Site of Data Collection	-0.01	0.06	-0.01	0.07	-0.01	0.06	0.04	0.04	0.07	0.04
Depression	0.03	0.06	0.03	0.07	0.02	0.07	0.06	0.04	0.06	0.05
Life Satisfaction	0.06	0.04	0.06	0.04	0.04	0.04	0.05*	0.02	0.03	0.03

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *$p < .05$. **$p < .01$.

Table S20

Truth and Bias Model Analyses for Snapchat Use with Depression and Life Satisfaction as Moderators of Tracking Accuracy and Directional Bias Across Different Model Specifications

	Model Type									
	Original Model		Heteroscedasticity-Corrected		Restricted-to-Active Use		Robust Regression		Influential Observations Removed	
Predictors	b	SE	b	SE	b	SE	b	SE	b	SE
Average Bias and Accuracy										
Average Directional Bias	28.26**	1.36	28.26**	1.45	25.71**	1.39	21.12**	0.89	23.20**	0.97
Average Tracking Accuracy	0.67**	0.04	0.67**	0.07	0.73**	0.04	0.70**	0.03	0.73**	0.04
Moderators of Directional Bias										
Site of Data Collection	1.89	1.38	1.89	1.46	0.01	1.41	1.22	0.91	1.05	0.98
Depression	4.56*	2.30	4.56	2.53	5.80*	2.32	3.30*	1.50	3.50*	1.70
Life Satisfaction	1.38	1.17	1.38	1.27	2.00	1.19	1.07	0.76	1.93*	0.84
Moderators of Tracking Accuracy										
Site of Data Collection	-0.03	0.04	-0.03	0.07	0.01	0.04	-0.01	0.03	-0.08*	0.04
Depression	0.11	0.07	0.11	0.13	0.08	0.07	0.03	0.05	0.12	0.07
Life Satisfaction	0.08*	0.04	0.08	0.07	0.07	0.04	0.01	0.02	0.04	0.03

Note. Site of Data Collection was effect-coded, such that -1 = the Southwestern University and 1 = the Midwestern university. All continuous predictors were grand-mean centered. Robust regression analyses used z-tests to determine significance level. *$p < .05$. **$p < .01$.