Tennessee v. Meta
Case No. 23-1364-IV
PX-00111

Document Provided in Native Format

social impact
well-being

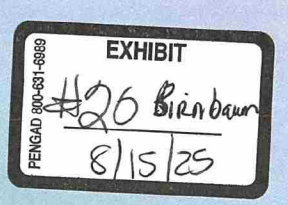
Problematic Use On-Platform Measurement

Concentration survey results for IG

Dec 2022- Jan 2023

Elena Davis and Josh [redacted]

X-Meta Well-being



TL;DR - IG

To advance our understanding and measurement of potential problematic use (PU) on our platforms, this workstream:

1) Developed and validated new survey measurement of users' perceived control of their use of Facebook and Instagram
2) Surveyed ~30K youth and ~30K adult users of IG in 12 countries
3) Identified on-platform behaviors associated with perceived low control

	Findings	Takeaways
1)	Users' perceived control can be validly measured across age cohorts with questions assessing **low attentiveness** around platform use (i.e., spending more time, checking, and consuming content without realizing it) and **low capability** to change platform use (i.e., low confidence in ability to reduce time, checking, or interactions)	These survey items can be adapted for use in product evaluations (designed on a case-specific basis) to understand perceived impact of products.
2)	Rates of perceived low control are similar for youth and adults, with low attentiveness (~34-40%) more common than low capability (13-20%)	While there are no objective cut-offs for defining 'low' control, we see that this is a general issue experienced by both youth and adults. Youth-first products can be explored to extend to gen pop.
3)	Perceived low control nearly fully explains the small association between on-platform behaviors and perceiving that one's platform use prevents fulfilling other responsibilities in life ('life interference').	Focusing on addressing areas on our platforms where users are likely to experience perceived low control is a viable strategy for making progress in this space.
4)	The ability to predict perceived control using logged platform behaviors is modest, but improved over past PU models: 0.77 AUC for attentiveness (precision/recall ~0.5/0.7) and 0.80 for capability (~0.4/0.4).	We have improved predictive precision by focusing on control, but on-platform behaviors likely cannot be used to directly measure (or proxy) perceived control.
5)	Global behaviors associated with perceived low control reflect high usage, evening + late night time spent, and specific behaviors, like no engagement repeat sessions and video chaining.	Our next step is to iterate with the team on appropriate cut-offs (e.g., p90 usage) and prioritized behaviors (e.g., evening + late night time) to use for heuristics and goaling, noting that these are *not* definitional of PU.
6)	Specific product usage associated with perceived low control includes content consumption surfaces (Explore, Reels Tab, Home) and active engagement (Stories creation, Direct).	Existing or novel control solutions may be provided in-context within these products, such as nudges within Reels or extensions of Quiet Mode to support limits and boundaries around active engagement.

Slide 2

1 Could you include the r^2?
Ashley [redacted] 2/10/2023

1

Approach

Slide 3

1 I think in your background slides you need to remind people about your PU mental model regarding how attention, awareness, capability, etc all fit together. I've forgotten it entirely so now I'm having trouble laddering up all the pieces again.
Judy [redacted] 2/6/2023

Context: Success for problematic use workstream

Overall success for PU is: *People have the tools that they need and are empowered to control their time and experiences on our platforms.*

Measurement will encompass: Signals of product efficacy + user value (e.g., behavior change metrics; product engagement; perceived impact) and positive perception (e.g., perceived usefulness; product awareness) *[to be aligned with measurement framework for all Youth + Well-being efforts across the company]*

Priority		Detail	Status
1)	**Meet youth industry practice requirements**	Ensure we are providing PU-related product solutions that are on par with the industry, based on regulatory expectations and competitor feature sets	IG feature set has largely met this bar, pending confirmation that products provide educational information in context (details TBD)
2)	**Measure success of products**	Ensure that the features are a) effective; b) bringing user value; c) reaching the relevant users	H1 2023 priority through behavior change goal, product evaluations and sizing, and product awareness research
3)	**Identify and close additional gaps**	Develop updated behavioral metrics to use in goaling Identify product usage and/or behaviors that are associated with perceived low control for additional investment	**This report** + next steps

Goal of this workstream: Identify on-platform behaviors that are associated* with user perception** of low control

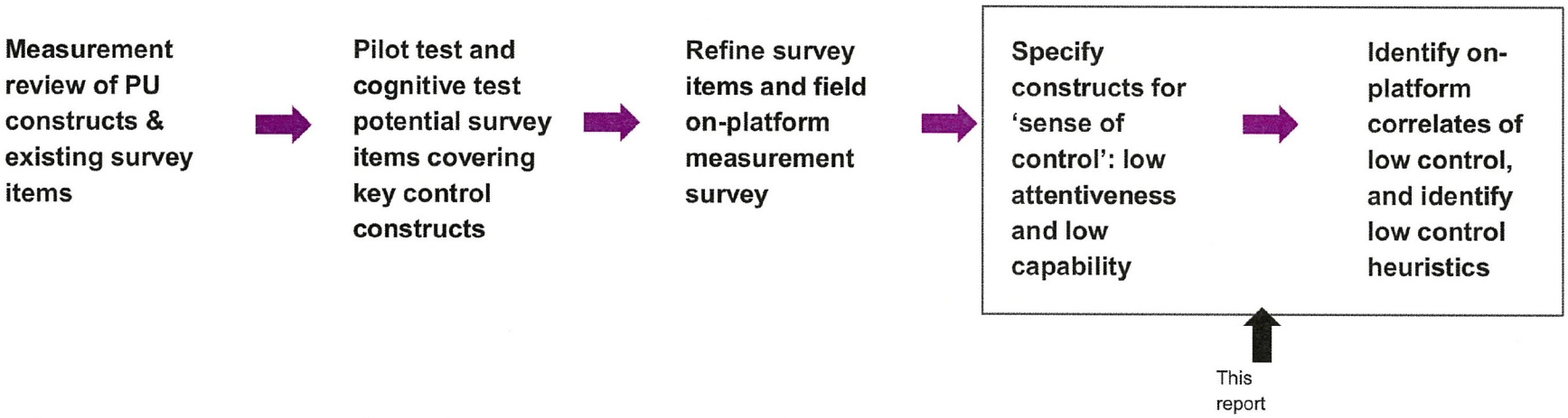
Outline of results

1) Measurement approach and rates of perceived low control
2) General behavior patterns (time spent, sessions, time of day) associated with perceived low control
3) Specific product usage associated with perceived low control

*All associations in this report are non-causal, i.e., there can be no conclusions drawn about the direction of how behavior and perception may impact each other.
**All user perception measures come from self-report surveys on the participants' subjective self-assessment of their experiences with sense of control and their perceptions of the platform.

Approach: Updated survey measurement of potential problematic use

Focus is on measuring users' perception of control over their time and experiences on our platforms (see note)



We identified two constructs of perceived control and measured them in an on-platform survey

Low attentiveness around use (possible unintended, unaware, and/or habitual use)	**Perceived control capability** (perceived self-efficacy to exert control, if desired)
1) How often do you find yourself spending more time than you wanted to without realizing it, when using Facebook/ Instagram?	1) How confident are you that you would be able to reduce the amount of time you spend using Facebook/ Instagram, if you wanted to?
2) How often do you find yourself checking Facebook/ Instagram out of habit, without realizing it? *Checking includes opening the app, checking a notification, or refreshing your feed for updates.*	2) How confident are you that you would be able to reduce how often you check Facebook/ Instagram, if you wanted to?
3) How often do you find yourself getting caught up viewing a lot of content on Facebook/ Instagram without realizing it?	3) How confident are you that you would be able to stop yourself from interacting with people you don't want to on Facebook/ Instagram? *Interacting includes responding to other people through comments or direct messages.*

Additional item on "life interference" adapted from the Main Tracking Survey was also included: How often does your use of Facebook/ Instagram prevent you from fulfilling your responsibilities in other areas of your life (for example, work, school, or relationships)?

These two areas of control are consistent with literature suggesting lack of control over social media use may come in two forms - habitual or low attention use; and lack of ability to control use (1, 2, 3). These also map to models of self-regulation of behavior that define two systems, an automatic one (akin to low attentiveness) and a conscious, goal-monitoring one (akin to exerting deliberate control) (4). See Appendix for details on survey development and item selection.

Details of survey sample, fielding, and analysis

Sample

N=30,699 adults and 30,346 youth (13-17) on Facebook
N=30,810 adults and 30,756 youth (13-17) on Instagram
Across 12 countries: AU, BR, EG, GB, KR*, ID, IN, MX, NG, PH, TR, US (**KR was IG only*)

Data collection

Sampling was stratified to be 50% users who exhibited 1 or more behaviors previously associated with PU (no engagement late night sessions, late night product switching, late night long sessions, no engagement repeat sessions, high frequency notification checking, or high volume of short sessions) and 50% users who exhibited none of those behaviors L7

Data was collected late November - early December 2022 (11/18-12/14 on FB; 11/21-12/8 on IG). Configs targeting less active users took longer to reach quotas.

Analysis

1 We verified that survey items formed 2 factors and created composite scores by averaging across the 3 attentiveness items and the 3 capability items.

Survey responses were weighted using Graviton weights to reflect the Facebook and Instagram user population across the 12 countries during the time frame of the survey, based on recommended indices from SWAMP: gender, country, age, indices of use (IG: followers, followings, creator, business, private account; FB: L30 + mobile and web, days since confirmed, friend count, subscription count profile picture), and device.

Associations between on-platform signals and survey measures were analyzed in 3 ways to confirm converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and simple correlation

Slide 8

2 Weighting question -- given that you used a stratified sample, do the two populations have the same distribution for the SWAMP characteristics (e.g. gender, country) as gen pop? If not, I don't know if the "out of the box" Graviton weights would be appropriate. I remember the team at some point was working to allow you to specify your population frame and create weights using that, but not sure whether those weights are currently available (/this plug and play option is available). As a short fix, though, if your results are the same with or without weights, that is something you can state.
Ashley [redacted] 2/14/2023

1 we did stratified sampling at data collection for high + low usage (which I think helped, the low usage segments took longer to collect so maybe we actually got a more representative sample), then combined the whole sample for weighting. interestingly the weighted data didn't adjust too much descriptives-wise, and I report out youth and adults segment separately.

country was definitely adjusted; I do have the read-outs from graviton and we could look together if you would be able to point out your specific Qs
Elena Davis, 2/14/2023

Results: Instagram

Summary of users' report of perceived low control

How common are experiences of perceived low control on Instagram?

Rates of reporting the *lowest* scores (<=2 on 5 pt composite scale; or 1 or more lowest responses on any item):

Low attentiveness: 38-39% youth, 34-40% adults
Low capability: 18-19% youth; 13-14% adults

How do these experiences differ across ages and markets?

Adults were slightly more likely to report low attentiveness; youth were slightly more likely to report low capability (small significant differences)

Markets differ: higher perceived control in western/high income (US, GB, AU); lower in some emerging markets (NG, EG, ID) and also relatively low in KR - possibly due to cross-cultural differences in perceptions of control and role of social media in people's lives

How does perceived low control relate to other perceptions of the app?

Users reporting low control also report higher 'life interference' from their app use (i.e., perceiving that use of Instagram prevents the fulfillment of other responsibilities). Low attentiveness: 20-25% higher life interference; low capability 17-23% higher life interference.

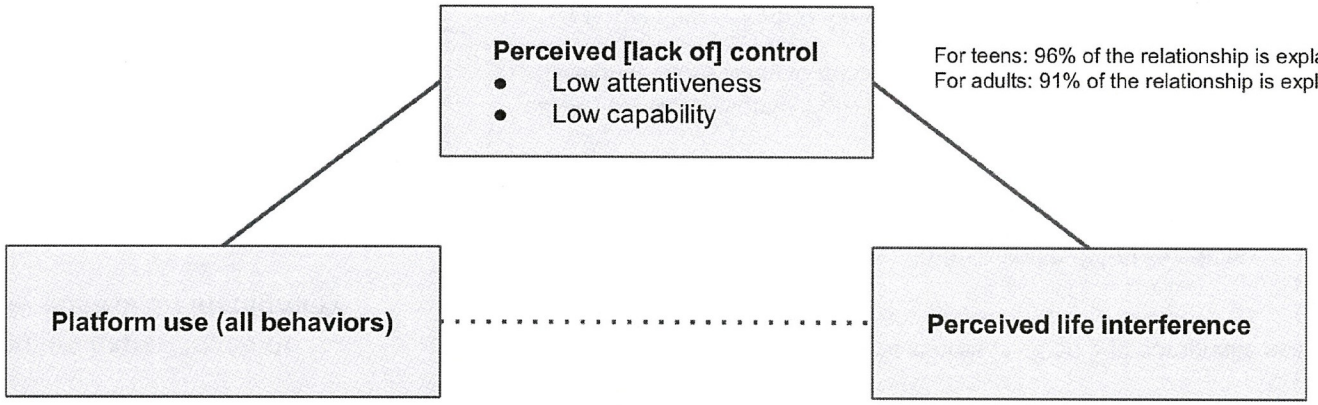
Low control is *not* strongly related to perceiving value from the app or views on whether the app is trustworthy or not (correlations < .14)

Descriptive details starting here

Perceived control explains how platform use may be related to perceived 'life interference'

Previous measurement work has included a focus on 'life interference' as an important outcome related to problematic use; i.e., perceiving that one's platform use is preventing the fulfillment of other life responsibilities.

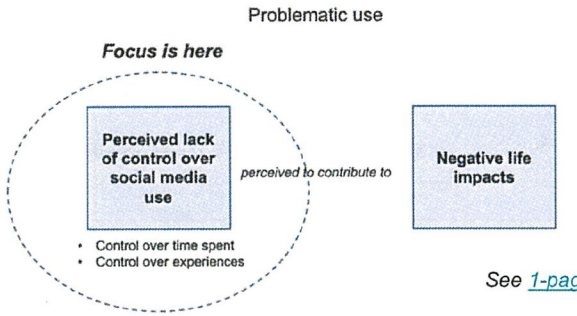
Here we see that perceived life interference is *almost entirely* related to actual platform use through perceived low control. We don't see an association between platform use and life interference except through perceived control.



Beyond the effects of perceived low control, platform signals only explain 0.7% of the variance in teen users reported life interference and 1.8% of the variance for adults.

Focusing on control increases our measurement precision

This supports our approach to focus on control: from both a conceptual and measurement standpoint, perceived control over platform use is more proximal to the actual on-platform experience, including behaviors that we can measure and potentially change.



See 1-pager here

Our overall models predicting perceived control are stronger than those that include life interference or other life impact measures, likely because there is less noise. Our updated heuristics will likely be better as well.

	Precision	Recall	AUC
Former PU model on FB Outcome: low control + negative life impact	0.17 all ages	0.73 all ages	0.66 all ages
Current PU model on IG Outcome: low attentiveness	0.48 youth 0.51 adults	0.70 youth 0.69 adults	0.77 youth 0.78 adults
Current PU model on IG Outcome: low capability	0.36 youth 0.50 adults	0.45 youth 0.36 adults	0.81 youth 0.80 adults

Specific behaviors associated with low perceived control on IG

Explanatory power for low attentiveness overall is 2-3x as strong as low capability. Predictors largely consistent across youth and adults.

Stack rank	Behavior	Attentiveness	Capability	Detail
1	**Total time spent**	x	x	Time spent and sessions (higher overall use) are top predictors
2	**Number of sessions**	x	x	
3	**Evening + early night time spent (7-10 pm; 10 pm-12 am)**	x	x	Whereas past PU research implicated late night time spent (1, 2), this study also identifies evening time spent as a predictor of perceived low control, slightly stronger than late night time spent.
4	**Passive evening + early night time spent (7-10 pm; 10 pm-12 am)**	x	x	Whereas past PU research found no association with sessions being more active or passive on average (2), certain no engagement sessions (repeat and late night) were found (1). Here we see passive sessions in evening and early night associated with perceived low control.
5	**Late night time spent (12am-4am) (overall and passive)**	x	x	Overall late night time spent, as well as passive time spent (slightly less strong), emerging here, consistent with past work.
6	**No engagement repeat sessions**	x	x	Consistent with past heuristic work (1) and possibly related to checking behavior. Relatively stronger as a predictor for adults
7	**Video chaining**	x		New finding, a behavior specifically associated with perceived low attentiveness. Particularly related to frequency of chaining, rather than chaining depth.
8	**Short sessions**		x	New finding, a behavior specifically associated with perceived low capability.

Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation

Specific product usage associated with low perceived control on IG

Stack rank	Product	Attentiveness (Youth and/or Adults)	Capability (Youth and/or Adults)	Detail
1	**Explore tab**	Y + A	Y + A	Total time spent and passive time spent
2	**Stories creation**	Y	Y + A	Frequency of stories creation
3	**Reels tab**	Y + A		Total time spent and passive time spent
4	**Direct**	Y	Y + A	Long sessions, total time spent
5	**Notifications**	Y + A	Y + A	Time spent, impressions, sessions originating from push
6	**Home**	Y + A	Y	Total time spent and passive time spent Note that *share* of time spent on Home was associated with *higher* attentiveness
7	**Profile**	A	Y	Total time spent and passive time spent

Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation

Takeaways

1) Many behaviors are associated with both forms of perceived low control, and predictors are largely similar for youth and adults.

1) Findings are consistent with the stance around focusing on late night use as important to support user control; extending to evening use may be a next step.

1) Expanding solutions to adult cohorts may also be beneficial.

1) Specific behaviors like video chaining, particularly on passive consumption surfaces like Reels, might be opportunities for attentiveness nudges or adjusting autoplay settings.

1) New product opportunities may exist in-context for specific surfaces or targeting specific behavior patterns - such as extending Quiet Mode to support limits and boundaries around active engagement.

Next steps

1) Focus this broad list to an abbreviated, non-overlapping list of 3-5 specific behaviors with relatively highest precision/recall to form updated PU heuristics

    - Leverage team priority area around late night use
    - Present alternatives and iterate on final set
    - Define thresholds

1) Evaluate sensitivity of heuristics: have we seen movement of these in experiments to date?

1) Opportunity size the potential impact of using these behaviors for additional targeting for Quiet Mode and Take a Break (e.g., expansion to evening hours)

    - Given a certain targeting threshold, what % of TAM is reached?
    - Update behavior change goals based on change in % of TAM reached and measured efficacy of the feature


16

social impact
well-being

Problematic Use On-Platform Measurement

Concentration survey results for FB

Dec 2022- Jan 2023

Elena Davis and Josh [redacted]

X-Meta Well-being

TL;DR - FB

To advance our understanding and measurement of potential problematic use (PU) on our platforms, this workstream:

1) Developed and validated new survey measurement of users' perceived control of their use of Facebook and Instagram
2) Surveyed ~30K youth and ~30K adult users of FB in 11 countries
3) Identified on-platform behaviors associated with perceived low control

	Findings	Takeaways
1)	Users' perceived control can be validly measured across age cohorts with questions assessing **low attentiveness** around platform use (i.e., spending more time, checking, and consuming content without realizing it) and **low capability** to change platform use (i.e., low confidence in ability to reduce time, checking, or interactions).	These survey items can be adapted for use in product evaluations (designed on a case-specific basis) to understand perceived impact of products.
2)	Rates of perceived low control are similar for youth and adults, with low attentiveness (~31-35%) more common than low capability (16-20%).	While there are no objective cut-offs for defining 'low' control, we see that this is a general issue experienced by both youth and adults. Youth-first products can be explored to extend to gen pop.
3)	Perceived low control nearly fully explains the small association between on-platform behaviors and perceiving that one's platform use prevents fulfilling other responsibilities in life ('life interference').	Focusing on addressing areas on our platforms where users are likely to experience perceived low control is a viable strategy for making progress in this space.
4)	The ability to predict perceived control using logged platform behaviors is modest, but improved over past PU models: ~0.7 AUC for attentiveness (precision/recall ~0.5/0.66) and ~0.75 for capability (~0.3/0.55).	We have improved predictive precision by focusing on control, but on-platform behaviors likely cannot be used to directly measure (or proxy) perceived control.
5)	Behaviors associated with perceived low control reflect high + frequent usage, response to notifications, evening + passive time spent, and specific behaviors like no engagement repeat sessions and surface switching.	Our next step is to iterate with the team on appropriate cut-offs (e.g., p90 usage) and prioritized behaviors (e.g., evening + passive time spent) to use for heuristics and goaling, noting that these are *not* definitional of PU.
6)	Specific product usage associated with perceived low control include content consumption + engagement (Stories, News Feed, Reels) as well as notifications.	Existing or novel control solutions may relate to these products, such as contextual nudges within Stories or Reels, or Quiet Mode future directions to support limits and boundaries around notifications.

TL;DR - FB

To advance our understanding and measurement of potential problematic use (PU) on our platforms, this workstream:

1) Developed and validated new survey measurement of users' perceived control of their use of Facebook and Instagram
2) Surveyed ~30K youth and ~30K adult users of FB in 11 countries
3) Identified on-platform behaviors associated with perceived low control

Approach

Context: Success for problematic use workstream

Overall success for PU is: *People have the tools that they need and are empowered to control their time and experiences on our platforms.*

Priority		Detail	Status
1)	**Meet youth industry practice requirements**	Ensure we are providing PU-related product solutions that are on par with the industry, based on regulatory expectations and competitor feature sets	Facebook feature set largely meeting this bar, with launch of Feed Pause. Additional gaps to be identified and aligned with CSI leadership
2)	**Measure success of products**	Ensure that the features are a) effective; b) bringing user value; c) reaching the relevant users	Overall measurement approach in progress for all of Youth & Well-being, encompassing signals of product efficacy + user value and positive perception
3)	**Refine measurement and identify additional gaps**	Develop updated behavioral metrics to use in goaling Identify product usage and/or behaviors that are associated with perceived low control for additional investment	**This report** + next steps

Goal of this workstream: Identify on-platform behaviors that are associated* with user perception** of low control

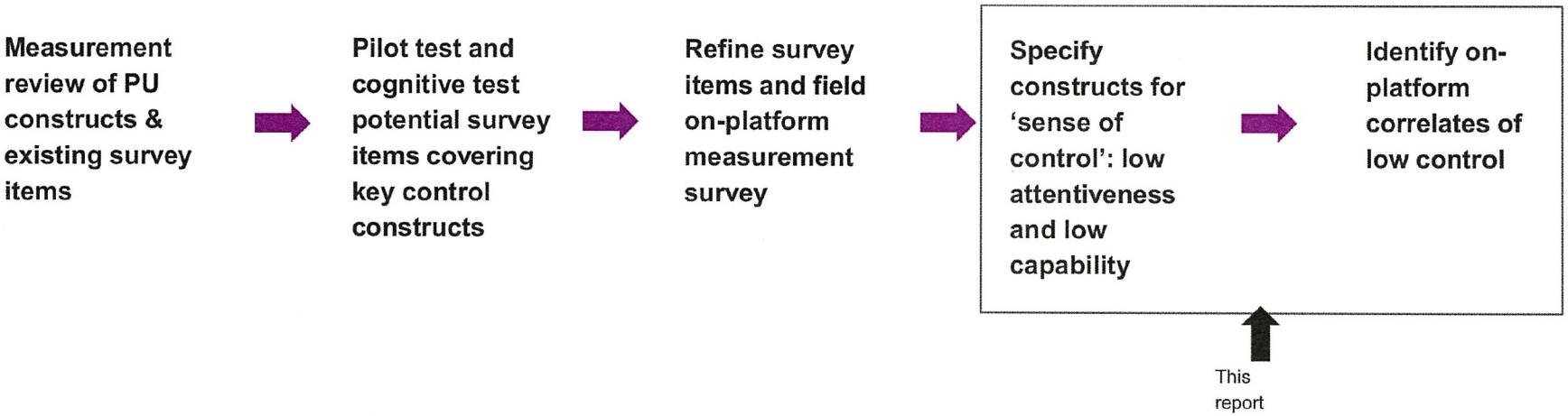
Outline of results

1) Measurement approach and rates of perceived low control
2) General behavior patterns (time spent, sessions, time of day) associated with perceived low control
3) Specific product usage associated with perceived low control

*All associations in this report are non-causal, i.e., there can be no conclusions drawn about the direction of how behavior and perception may impact each other.
**All user perception measures come from self-report surveys on the participants' subjective self-assessment of their experiences with sense of control and their perceptions of the platform.

Approach: Updated survey measurement of potential problematic use

Focus is on measuring users' perception of control over their time and experiences on our platforms (see note)



We identified two constructs of perceived control and measured them in an on-platform survey

Low attentiveness around use (possible unintended, unaware, and/or habitual use)	**Perceived control capability** (perceived self-efficacy to exert control, if desired)
1) How often do you find yourself spending more time than you wanted to without realizing it, when using Facebook/ Instagram?	1) How confident are you that you would be able to reduce the amount of time you spend using Facebook/ Instagram, if you wanted to?
2) How often do you find yourself checking Facebook/ Instagram out of habit, without realizing it? *Checking includes opening the app, checking a notification, or refreshing your feed for updates.*	2) How confident are you that you would be able to reduce how often you check Facebook/ Instagram, if you wanted to?
3) How often do you find yourself getting caught up viewing a lot of content on Facebook/ Instagram without realizing it?	3) How confident are you that you would be able to stop yourself from interacting with people you don't want to on Facebook/ Instagram? *Interacting includes responding to other people through comments or direct messages.*

Additional item on "life interference" adapted from the Main Tracking Survey was also included: How often does your use of Facebook/ Instagram prevent you from fulfilling your responsibilities in other areas of your life (for example, work, school, or relationships)?

These two areas of control are consistent with literature suggesting lack of control over social media use may come in two forms - habitual or low attention use; and lack of ability to control use (1, 2, 3). These also map to models of self-regulation of behavior that define two systems, an automatic one (akin to low attentiveness) and a conscious, goal-monitoring one (akin to exerting deliberate control) (4). See Appendix for details on survey development and item selection.

Details of survey sample, fielding, and analysis

Sample

N=30,699 adults and 30,346 youth (13-17) on Facebook
N=30,810 adults and 30,756 youth (13-17) on Instagram
Across 12 countries: AU, BR, EG, GB, KR*, ID, IN, MX, NG, PH, TR, US (**KR was IG only*)

Data collection

Sampling was stratified to be 50% users who exhibited 1 or more behaviors previously associated with PU (no engagement late night sessions, late night product switching, late night long sessions, no engagement repeat sessions, high frequency notification checking, or high volume of short sessions) and 50% users who exhibited none of those behaviors L7

Data was collected late November - early December 2022. Configs targeting less active users took longer to reach quotas.

Analysis

We verified that survey items formed 2 factors and created composite scores by averaging across the 3 attentiveness items and the 3 capability items.

Survey responses were weighted using Graviton weights to reflect the Facebook and Instagram user population across the target countries during the time frame of the survey, based on recommended indices from SWAMP: gender, country, age, indices of use (IG: followers, followings, creator, business, private account; FB: L30 + mobile and web, days since confirmed, friend count, subscription count profile picture), and device.

Associations between on-platform signals and survey measures were analyzed in 3 ways to confirm converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and simple correlation

Results: Facebook

Summary of users' report of perceived low control

How common are experiences of perceived low control on Facebook?

Rates of reporting the *lowest* scores (<=2 on 5 pt composite scale; or 1 or more lowest responses on any item):

Low attentiveness: 31-34% youth, 32-35% adults
Low capability: 19-20% youth; 16-17% adults

How do these experiences differ across ages and markets?

Adults were slightly more likely to report low attentiveness; youth were slightly more likely to report low capability (small significant differences)

Markets differ: higher perceived control in western/high income markets (US, GB, AU); lower in emerging markets (PH, NG, EG) - possibly due to cross-cultural differences in perceptions of control and role of Facebook in people's lives

How does perceived low control relate to other perceptions of the app?

Users reporting low control also report higher 'life interference' from their app use (i.e., perceiving that use of Facebook prevents the fulfillment of other responsibilities). Low attentiveness: 23-25% higher life interference; low capability 23% higher life interference.

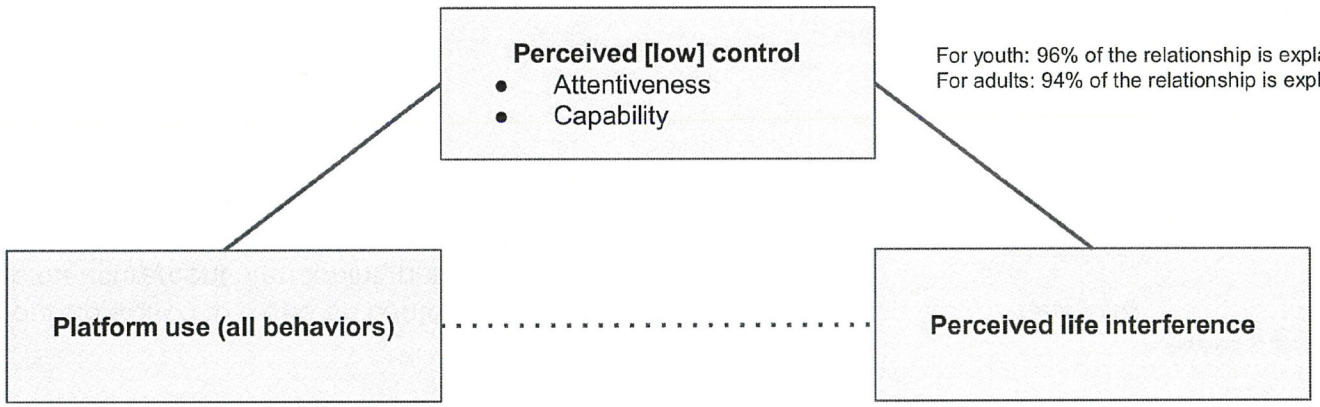
Low control is *not* strongly related to perceiving value from the app or views on whether the app is trustworthy or not (correlations < .15)

Descriptive details here

Perceived control explains how platform use may be related to perceived 'life interference'

Previous measurement work has included a focus on 'life interference' as an important outcome related to problematic use; i.e., perceiving that one's platform use is preventing the fulfillment of other life responsibilities.

Here we see that perceived life interference is *almost entirely* related to actual platform use through perceived low control. We don't see an association between platform use and life interference except through perceived control.



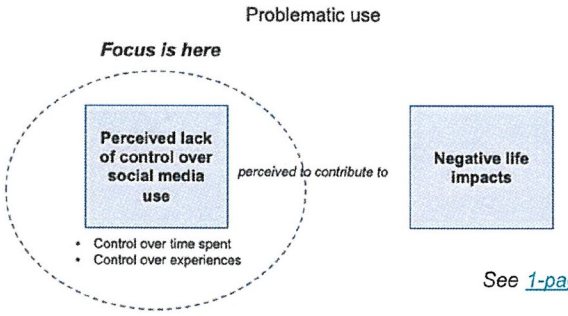
For youth: 96% of the relationship is explained by perceived control
For adults: 94% of the relationship is explained by perceived control

Beyond the effects of perceived low control, platform signals only explain 0.6% of the variance in youth reported life interference and 0.8% of the variance for adults.

Models were run by regressing life interference scores on control constructs, with the portions of life interference not mediated by the control constructs then being regressed on logged on-platform behaviors.

Focusing on control increases our measurement precision

This supports our approach to focus on control: from both a conceptual and measurement standpoint, perceived control over platform use is more proximal to the actual on-platform experience, including behaviors that we can measure and potentially change.



See 1-pager here

Our overall models predicting perceived control are stronger than those that include life interference or other life impact measures, likely because there is less noise. Our updated heuristics will likely be better as well.

	Baseline	Precision	Recall	AUC
Former PU model on FB Outcome: low control + negative life impact	~0.11	0.17 all ages	0.73 all ages	0.66 all ages
Current PU model on FB Outcome: low attentiveness	~0.3-0.35	0.54 youth 0.49 adults	0.67 youth 0.66 adults	0.74 youth 0.69 adults
Current PU model on FB Outcome: low capability	~0.15-0.2	0.30 youth 0.24 adults	0.59 youth 0.53 adults	0.76 youth 0.73 adults

Top global behaviors associated with low perceived control on FB

Explanatory power for low attentiveness overall is ~5x as strong as low capability. Predictors were largely the same for youth + adults.

Stack rank	Behavior	Detail	Attentiveness	Capability
1	**Total time spent**	Time spent is the top predictor for low control, stronger for attentiveness but relatively strong for both	x	x
2	**Sessions originating from push notification**	Consistent with past heuristics and PU research pointing to the role of responding to notifications (1, 2). Relatively important predictor for low capability.		x
3	**Number of sessions**	Note that number of sessions are highly correlated with total time spent, rs .65-.7. Consistent with past work (1, 2), session *length* was not a top predictor.	x	x
4	**Evening time spent (4-7pm; 7-10pm; 10pm-12am)**	Whereas past PU research implicated late night time spent (1, 2), this study identifies evening time spent as a stronger predictor of low control, particularly of attentiveness	x	
5	**Passive sessions (overall; evening 4-7pm, 7-10pm)**	Whereas past PU research found no association with active vs. passive sessions overall (2), certain no engagement sessions (repeat and late night) were found (1). Here we see passive sessions overall and evening-specific associated with low control.	x	x
6	**Surface switching**	Consistent with past heuristic work (1). Relatively important predictor for low attentiveness	x	
7	**No engagement repeat sessions**	Consistent with past heuristic work (1). Relatively important predictor for low capability.		x
8	**Late night sessions (overall + passive)**	Consistent with past heuristic work (1). Relatively important predictor for low capability.		x

Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation

Specific product usage associated with low perceived control on FB

Stack rank	Product	Low attentiveness		Low capability	
		Youth	Adults	Youth	Adults
1	**Stories**	Engagements (reactions, comments, and reshares) Engagement rate (per time spent)	Total time spent	Total time spent Engagements (reactions, comments, and reshares)	Total time spent Engagements (reactions, comments, and reshares)
2	**Notifications**	Sessions originating from push Notification checking within app	Sessions originating from push Notification checking within app	Sessions originating from push Shorter time to start session following push	Sessions originating from push
3 or 4	**News Feed**		Total time spent Passive time spent		
3 or 4	**Reels**	Total time spent Share of time spent	Total time spent Share of time spent Passive time spent Engagements (reactions, comments, and reshares)		
5	**Comments**	Total time spent Engagement (reactions, replies)	Total time spent Engagement (reactions, replies)	Total time spent Engagement (reactions, replies)	

Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation

Specific product usage associated with high perceived control on FB

Stack rank	Product	High attentiveness		High capability	
		Youth	Adults	Youth	Adults
1	**Photos**	Total time spent Share of time spent	Total time spent Share of time spent	Total time spent Share of time spent	Total time spent Share of time spent
2	**News tab** **low overall usage in sample (<2% WAU) so highly self-selected group*	Total time spent Engagement	Total time spent Engagement	Total time spent Engagement	Total time spent Engagement
3	**Groups**				Total time spent

Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation

Takeaways

1) Rates of perceived low control and on-platform correlates were largely similar for youth and adults. Solutions for both cohorts may be beneficial.

1) Global behaviors associated with perceived low control are around frequent use, responding to notifications, and cumulative high use, rather than longer individual sessions.

1) Passive time spent in the evenings and surface-switching behavior may be opportunities to offer attentiveness-type control features, with relevant products including News Feed and Reels.

1) Stories and notifications may be opportunities to explore additional control solutions, such as contextual nudges within Stories or Quiet Mode future directions to support limits and boundaries around notifications.

Next steps

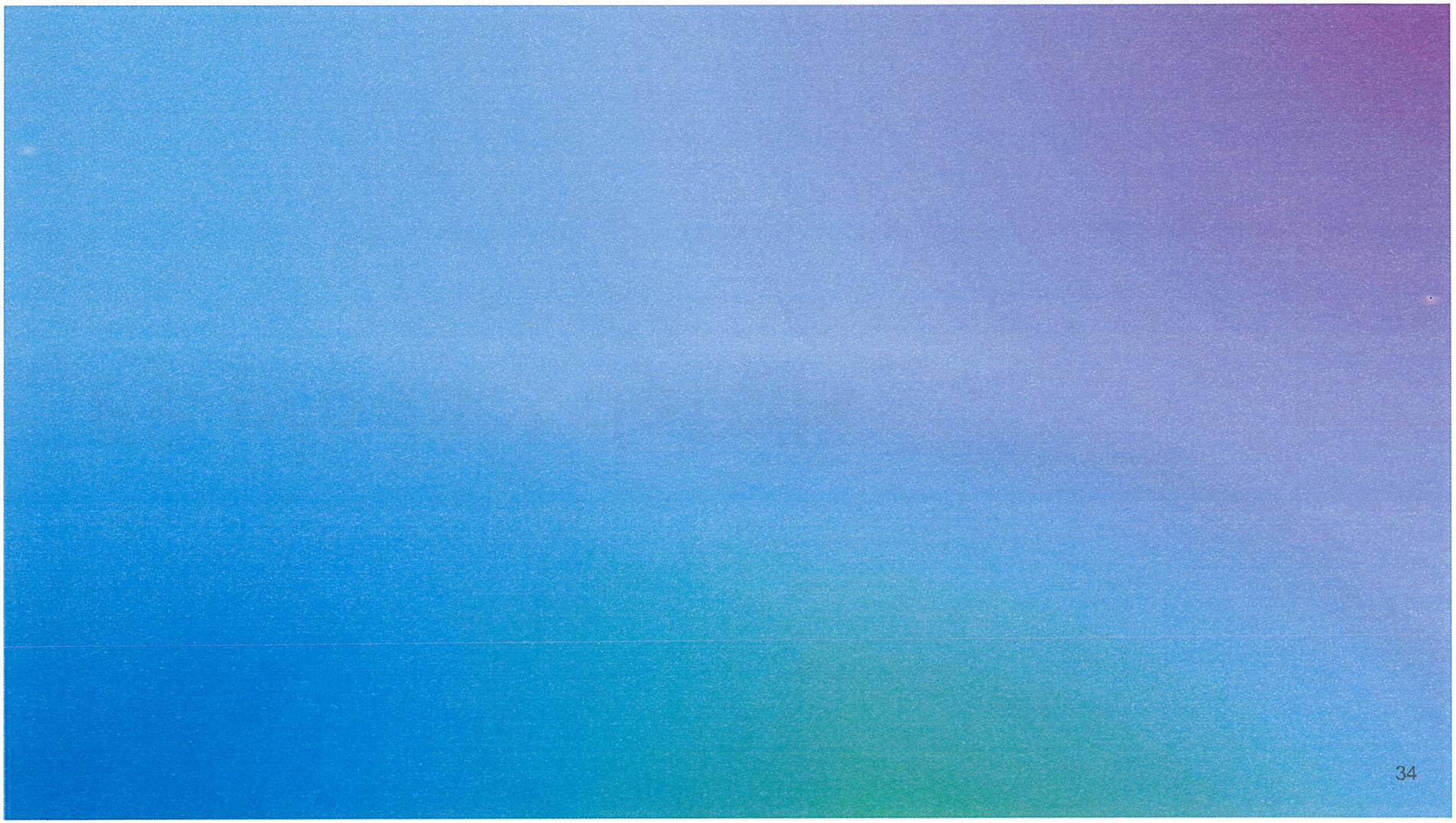
1) Focus the broad list of behaviors to an abbreviated, non-overlapping list of 3-5 specific behaviors with relatively highest precision/recall to form updated PU heuristics

    - Leverage team priorities and input
    - Present alternatives and iterate on final set
    - Define thresholds

1) Evaluate sensitivity of heuristics: have we seen movement of these in experiments to date?

1) Opportunity size the potential impact of using these behaviors for additional targeting of existing products

    - Given a certain targeting threshold, what % of TAM is reached?
    - Update product goals based on change in % of TAM reached and measured efficacy of the feature


34

FB product details

FB product usage associated with *lower* perceived control *(non-causal)*

2 Product	Low Attentiveness *Greater usage of these products associated* with more frequently spending more time, checking more, and consuming more content than intended, without realizing it*		Low Capability *Greater usage of these products associated* with perceived inability to reduce or limit time spent, checking, and interactions, even if wanting to*	
	Teens	**Adults**	**Teens**	**Adults**
Stories **Tier 1 predictor**	Engagements (reactions, comments, and reshares) Engagement rate (per time spent)	Total time spent	Total time spent Engagements (reactions, comments, and reshares)	Total time spent Engagements (reactions, comments, and reshares)
News Feed **Tier 1 predictor**		Total time spent Passive time spent		
Reels **Tier 3 predictor**	Total time spent Share of time spent	Total time spent Share of time spent Passive time spent Engagements (reactions, comments, and reshares)		

3 *Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation
** Tier of strength prediction is relative: Tier 1 = top 3 strongest predictor from regression and/or decision tree; Tier 2 = top 3 strongest predictor from one cohort or one form of low control; Tier 3 = top 10 predictor

Slide 36

2 highlighting these so we can return to them
Elena Davis, 1/25/2023

3 the definitions need to be more specifically operationalized and confirmed across all 3 analyses and cohorts; tiers are relative at present and these are general criteria
Elena Davis, 2/3/2023

FB product usage associated with *lower* perceived control *(non-causal)*

Product	Low Attentiveness *Greater usage of these products associated* with more frequently spending more time, checking more, and consuming more content than intended, without realizing it*		Low Capability *Greater usage of these products associated* with perceived inability to reduce or limit time spent, checking, and interactions, even if wanting to*	
	Teens	**Adults**	**Teens**	**Adults**
Notifications **Tier 1 predictor**	Sessions originating from push Notification checking within app	Sessions originating from push Notification checking within app	Sessions originating from push Shorter time to start session following push	Sessions originating from push
Comments (reading and reacting) **Tier 3 predictor**	Total time spent Engagement (reactions, replies)	Total time spent Engagement (reactions, replies)	Total time spent Engagement (reactions, replies)	

*Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation
** Tier of strength prediction is relative: Tier 1 = top 3 strongest predictor from regression and/or decision tree; Tier 2 = top 3 strongest predictor from one cohort or one form of low control; Tier 3 = top 10 predictor

FB product usage associated with *higher* perceived control *(non-causal)*

Product	High Attentiveness *Greater usage of these products associated* with <u>less frequently</u> spending more time, checking more, and consuming more content than intended, without realizing it*		High Capability *Greater usage of these products associated* with <u>higher perceived ability</u> to reduce or limit time spent, checking, and interactions, if wanting to*	
	Teens	**Adults**	**Teens**	**Adults**
Photos **Tier 1 predictor**	Total time spent Share of time spent	Total time spent Share of time spent	Total time spent Share of time spent	Total time spent Share of time spent
News tab **Tier 3 predictor** **low overall usage in sample (<2% WAU) so highly self-selected group*	Total time spent Engagement	Total time spent Engagement	Total time spent Engagement	Total time spent Engagement
Groups **Tier 3 predictor**				Total time spent

*Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation
** Tier of strength prediction is relative: Tier 1 = top 3 strongest predictor from regression and/or decision tree; Tier 2 = top 3 strongest predictor from one cohort or one form of low control; Tier 3 = top 10 predictor

Takeaways - Global behaviors

Overall lower control	Both forms of perceived low control are generally associated with higher use: • higher overall time spent • higher numbers of sessions • passive use *→ Longer time spent <u>per session</u> was not associated; the behavior is about frequent use and cumulative high use, rather than longer individual sessions. Passive sessions are the majority of sessions overall across users.*
Low attentiveness	Particularly characterized by: • more evening use • surface switching *→ This paints a picture of losing track of time when using the app at the end of the day (when self-regulation may be lower), and checking within the app which may be less intentional (because it is fast, it might be about checking, clearing jewels, or clicking around without deliberation)*
Low capability	Particularly characterized by: • sessions originating from push notifications • no engagement repeat sessions (returning to use the app within 10 minutes of the last session, with no engagement) • late night use *→ This paints a picture of users who are responsive to notifications and frequently return to the app (e.g., to check), and late night use that is possibly difficult to reduce (though we don't know why).*

Takeaways - Specific products

2 **Content consumption** Stories News Feed Reels	Stories usage is the strongest and most consistent predictor of low control. This may be due to content consumption affordances (like autoplay) but also about ephemerality, which may contribute to not wanting to miss updates and desire to respond to others in a timely way (consistent with social pressures, FOMO; expert input supports these hypotheses 1, 2; user research 1). Solutions that increase attentiveness (such as break nudges like Feed Pause) may be a particularly good fit for News Feed and Reels.
Responsiveness Notifications Comments	Notifications have been specifically pointed out as a possible trigger of or contributor to low control usage (1, 2, 3, 4); this may not only be due to initiating new sessions, but diversion within a session. Comment engagement is [?a mixed bag of value]
Specific uses Photos Groups News	Time spent on these types of surfaces, which require specifically navigating to them, may possibly indicate a more intentional behavior within app usage. News in particular is a low-usage surface, so this association is likely due to a self-selected sample (high news consumers + high perceptions of control)

Slide 40

1 Comments are in MSI so reactions and comments themselves flow into at least the MSI portion of ranking.

"his may not only be due to initiating new sessions, but diversion within a session" - I can vet this hypothesis. If we see that controlling for notifications there is little movement in the comments association it suggests comments -> notifs -> sessions isn't the mechanism.

I would be looking at both the low control association and the sessions association.
Josh [redacted], 1/31/2023

4 maybe connect w Tom about this? how comments are/aren't used in different ranking signals; their role in civic?
Elena Davis, 1/31/2023

2 Comments have a 0.5 correlation with time spent and surface switching and 0.3 with sessions.

When looking at the relationship between comments and low control usage we see that controlling for sessions and notifications (85% of the lift in sessions associated with comments is explained by notifications) ~50% of the effect size still remains for commenting behavior.

This remaining 50% is the diversion within the session - more surface switching and more time spent.

It appears roughly half of the relationship between low control and commenting has to do with getting diverted within sessions and the other half with increasing the sessions/checking.
Josh [redacted] 1/31/2023

IG product details

Theme: Continuous content consumption

Specific product usage associated with perceived *low* control

Product	Low Attentiveness *Greater usage of these products associated* with more frequently spending more time, checking more, and consuming more content than intended, without realizing it*		Low Capability *Greater usage of these products associated* with perceived inability to reduce or limit time spent, checking, and interactions, even if wanting to*	
	Teens	**Adults**	**Teens**	**Adults**
Explore tab* **Tier 1 predictor**	Total time spent Passive time spent	Total time spent Passive time spent	Total time spent Share of time spent Passive time spent	Total time spent Share of time spent
Reels tab **Tier 2 (?or 3) predictor**	Total time spent Passive time spent	Total time spent Passive time spent		
Video chaining (any surface)** **Tier 2 (?or 3) predictor**	Total time spent Frequency of chaining	Total time spent Frequency of chaining		Total time spent Frequency of chaining
Home **Tier 3 predictor**	Total time spent Passive time spent >1 hour/week	Total time spent Passive time spent >1 hour/week	Passive time spent	

*Explore tab also includes behaviors like tapping into full screen + chaining.

**Frequency of video chaining *at any depth* is associated with perceived low control. Frequently getting a few videos deep is twice as important in association than daily chaining depth (i.e., frequently getting a few videos deep matters more than occasionally getting 10-20+ videos deep).

*Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation
** Tier of strength prediction is relative: Tier 1 = top 3 strongest predictor from regression and/or decision tree; Tier 2 = top 3 strongest predictor from one cohort or one form of low control; Tier 3 = top 10 predictor

Slide 42

6 @[redacted]@meta.com you had a previous point that 'Across all three surfaces, chaining depth is a specific factor." but we are saying that chaining frequency at any depth is more important... can you help reconcile?
Elena Davis, 2/2/2023

3 It's the frequency of getting into chaining that is important, overall impressions/video seen in chaining is important however the max depth a user gets into chaining is much more weakly related.

Meaning it isn't about how far you go down the chaining here but the repeatedly getting into chaining even if you only go a few videos deep.

My hunch is this has to do with explore, clips tab, and when you click into full screen on video off of feed, people may get caught up for a couple videos or until they hit one they aren't interested in, then they reset to whatever surface but get right back into chaining.
Josh[redacted], 2/2/2023

8 Tom asking (how well does each of these product experiences independently predict the outcome? correlation)
Elena Davis, 2/2/2023

7 noting to clean up this operationalization and confirm for all products
Elena Davis, 2/6/2023

3 This is kind of weird. Why not just use the features that showed up as highest weights across both models? It just makes this slide with the tiering structure confusing to interpret.

@[redacted]@fb.com
Judy[redacted]2/6/2023

4 I wonder if this is all better presented in a bunch of figures and use a small table at the very end to summarize?
Judy[redacted], 2/6/2023

2 Can you match these cell description with the names you used for slide 12?
Judy[redacted] 2/7/2023

5 ah so they don't! will be good to clarify
Elena Davis, 2/7/2023

9 noting to make sure I am capturing this correctly
Elena Davis, 2/8/2023

4 This one for adults is similar to the attention result where it's both total time spent and the frequency of getting into chaining.
Josh[redacted], 2/8/2023

Specific product usage associated with perceived *low* control

11 oduct	Low Attentiveness *Greater usage of these products associated* with more frequently spending more time, checking more, and consuming more content than intended, without realizing it*		Low Capability *Greater usage of these products associated* with perceived inability to reduce or limit time spent, checking, and interactions, even if wanting to*	
	Teens	**Adults**	**Teens**	**Adults**
Stories creation **Tier 1 predictor**	Frequency		Frequency	Frequency
Direct*** **Tier 3 predictor**	Long sessions		Total time spent	Total time spent
Notifications **Tier 3 predictor**	Time spent on activity Notification impressions Sessions originating from push	Time spent on activity Notification impressions	Time spent on activity Sessions originating from push	Time spent on activity Sessions originating from push

***Direct time spent still a predictor even controlling for notifications.

*Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation
** Tier of strength prediction is relative: Tier 1 = top 3 strongest predictor from regression and/or decision tree; Tier 2 = top 3 strongest predictor from one cohort or one form of low control; Tier 3 = top 10 predictor

Slide 43

10 yeah so these are specific to notifications -- slide 12 is about behavior patterns across surfaces, these findings are about specific products
Elena Davis, 2/7/2023

6 Got it. This seems quite complex, no? Are you then saying that each team should be goaling on a different heuristic depending on their product type
Judy [redacted], 2/7/2023

5 Wait where did these come from? They're not from slide 12
Judy [redacted], 2/7/2023

11 ah so -- this is more to guide the product intervention, like which surfaces might we need to layer existing products in or develop new product changes?
Elena Davis, 2/7/2023

Specific product usage associated with perceived *low* control

Product	Low Attentiveness *Greater usage of these products associated* with more frequently spending more time, checking more, and consuming more content than intended, without realizing it*		Low Capability *Greater usage of these products associated* with perceived inability to reduce or limit time spent, checking, and interactions, even if wanting to*	
	Teens	**Adults**	**Teens**	**Adults**
Profile		Total time spent Passive time spent	Total time spent Engagement	

*Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation
** Tier of strength prediction is relative: Tier 1 = top 3 strongest predictor from regression and/or decision tree; Tier 2 = top 3 strongest predictor from one cohort or one form of low control; Tier 3 = top 10 predictor

Specific product usage associated with perceived *high* control

Product	High Attentiveness *Greater usage of these products associated* with <u>less frequently</u> spending more time, checking more, and consuming more content than intended, without realizing it*		High Capability *Greater usage of these products associated* with <u>higher perceived ability</u> to reduce or limit time spent, checking, and interactions, if wanting to*	
	Teens	**Adults**	**Teens**	**Adults**
Home	Share of time spent			

*Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation
** Tier of strength prediction is relative: Tier 1 = top 3 strongest predictor from regression and/or decision tree; Tier 2 = top 3 strongest predictor from one cohort or one form of low control; Tier 3 = top 10 predictor

Usage of products associated with perceived low control

Continuous content consumption Explore tab Reels tab ~~Video chaining~~ Home esp long passive	Content consumption in Explore, Reels, and Home - especially passive - is an experience within which we can offer control features. For example, Take a Break with more nuanced time spent targeting. Frequently viewing chained content, regardless of which surface, is associated particularly with low attentiveness. This makes sense as a product affordance that may contribute to people spending time or consuming content without realizing it (expert report). Importantly, *depth* of chaining is less associated. This would support enacting more limits on autoplay of next videos that occurs across surfaces, and/or adding pauses or friction. Occasional longer immersive viewing sessions may be more intentional especially if they are paired with features that support awareness and limits around that experience.
Active engagement Stories creation Direct Notifications	Stories creation is the strongest predictor of perceived low capability to reduce use, and also associated with low attentiveness - though has been found to be a high intent behavior (1). Content creation was also identified in previous PU research (2), as highly active users are also more likely to create content. Creation likely leads to more responses and thus likelihood of returning to use the app, all in a time-bound ephemeral way. Direct messages have been previously found associated with PU heuristics and related survey outcomes, but also are related to high session satisfaction (3). Control opportunities here will have to find a balance. Notifications have been specifically pointed out as a possible trigger of or contributor to low control usage (4, 5, 6, 7); this may not only be due to initiating new sessions, but diversion within a session.
Other Profile Home share of time	Profile time spent was also identified in previous PU research (1), but unclear why. Past work found no differences in browsing own vs. other profiles (2). Profile use may be related to revisiting own past content; social curation, editing/updating (3; 4; 5). Low control experiences especially with passive profile time spent could be a particular form of content consumption. Share of time spent on Home associated with *higher* control might be due to trading off with time spent on other surfaces associated with low control.

Slide 46

13 not sure if I should speculate further here about possible cohorts who spend time on Home almost exclusively compared to other uses of IG, or the connected/unconnected distinction
Elena Davis, 2/3/2023

7 I think this is a bit of a strange bucket. Makes me wonder if this slide and trying to bucket by product type is even necessary.
Judy [redacted], 2/6/2023

12 maybe move this up to the global behaviors section
Elena Davis, 2/8/2023

5 Yeah that's interesting, it's a surface agnostic thing. It's really about are you getting into chaining from whatever entry point with high frequency
Josh [redacted], 2/8/2023

8 Other ways to label this: "Passive feed consumption"?
Judy [redacted], 2/8/2023

6 Explore (it's video originating from explore not like on the grid surface itself) and reels tab, maybe are a little different than what feed consumption invokes.

Since they are full screen video experiences that will transition into the next video upon completion of the first if you let it.
Josh [redacted] 2/8/2023

How well can we predict user self-reported perceived control using on-platform signals?

Overall our predictive power for perceived low control is modest: precision of ~0.5 and recall of ~0.7 for attentiveness and ~0.4 on both precision and recall for capability. We can predict ~30% of the variance in capability and 37% of the variance in attention.

15 10	Attentiveness			Capability		
	Precision	Recall	AUC	Precision	Recall	AUC
All data - Youth	.48	.70	0.766	.36	.45	0.807
All data - Adults	.51	.69	0.775	.50	.36	0.801
Top 10 - Youth	.40	.71	0.696	.23	.33	0.693
Top 10 - Adults	.42	.71	0.695	.20	.29	0.674

Slide 47

7 AUC was 0.67 for the previous work and we are ~0.7 for life interference this time
Josh [redacted], 2/6/2023

8 Capability AUC is better because those precisions are pretty large increases vs. the baseline rate.
Josh [redacted], 2/6/2023

14 wonder how we can put some of these numbers in context from carlos's past work, seems like we are doing better for attentiveness; not sure how to interpret his AUC?
Elena Davis, 2/6/2023

15 I'm tempted to add another slide capturing past work, and situating this work as a step forward
Elena Davis, 2/6/2023

16 huh but not better for control?

why would that be
Elena Davis, 2/7/2023

9 For reference the simplified models taking the top 10 behaviors are doing better than the full versions on life interference
Josh [redacted], 2/7/2023

10 Sorry, I think I wrote this confusingly.

The simplified models for low control are more accurately predicting than what we get from the full model for life interference.

This is just a supporting point about being more able to predict low control or it having less noise than life interference.

While not strictly true because of potential overfitting concerns, having a larger set of features should result in equal or better prediction.

If we look at a simplified version of life interference it will be close but below the full model (IIRC it was something like 0.64-0.65 AUC for the simplified model compared with 0.67-0.7 for full model - Carlos also found something similar).
Josh [redacted], 2/7/2023

Global behavior correlations - FB youth

Correlations	Attentiveness	Capability
surface_visits	-0.12	-0.05
engagements	-0.06	-0.04
time_spent	-0.14	-0.06
push_notif_session	-0.14	-0.07
time_to_sess_from_push	0.06	0.05
late_night_ts	-0.10	-0.02
ts_4_7	-0.12	-0.05
ts_7_10	-0.12	-0.05
ts_10_12	-0.11	-0.04
late_night_sessions	-0.12	-0.04
no_eng_sessions	-0.11	-0.04
late_night_high_surface_switching_sessions	-0.09	-0.03
notif_checks	-0.10	-0.02
short_sessions	-0.09	-0.03
repeat_no_eng_session	-0.12	-0.06
late_night_no_eng_sessions	-0.09	-0.03
no_engagement_4_7	-0.10	-0.04
no_engagement_7_10	-0.10	-0.04
no_engagement_10_12	-0.10	-0.03

Boosted Decision Tree for Attentiveness - FB youth

Feature (Top 10)	Gain (marginal effect of predictive accuracy)	Frequency (of behavior)	Correlation
photo_ts	0.154772658	0.060150376	-0.08
notif_ts_share	0.0964493844	0.033834586	-0.07
nf_ts_share	0.0652362876	0.033834586	-0.11
story_ts_share	0.0570405409	0.033834586	-0.06
feedback_ts_share	0.0431154843	0.026315789	-0.06
notif_checks	0.0407811896	0.056390977	-0.10
photo_ts_share	0.0402052754	0.030075188	
feedback_ts	0.0396432915	0.022556391	-0.08
ts_7_10	0.0290890637	0.037593985	-0.12
nf_ts	0.0282388637	0.041353383	-0.12
reels_ts_share	0.0270582468	0.026315789	-0.09
engagements	0.0259578992	0.026315789	-0.06
ts_4_7	0.0244354999	0.022556391	-0.12
push_notif_session	0.0221486385	0.026315789	-0.14
ts_10_12	0.018119968	0.030075188	-0.11
feedback_engagement_per_min	0.0175495671	0.030075188	-0.03
no_engagement_4_7	0.0161219838	0.022556391	-0.10

Overall behavior patterns - Facebook

havioral signal	Low attention Strength of associations		Low capability Strength of associations		
	Teens	Adults	Teens	Adults	Overall predictor strength for linking on-platform signals with low control are 5x stronger for predicting low attention than low capability; prediction for low capability overall low with push sessions + no engagement repeat sessions on top
Total time spent	~6%	~6%	~1%	N/A (not top 10 predictor)	
Sessions originating from push notification	~5.5%	~5.75%	~1%	~1%	Push notifications more salient as a predictor for FB compared to IG Interesting that this is right up there with number of sessions- could sessions originating from a push be overall less intentional sessions?
Sessions (number)	~5%	~6%	~.75%	~.5%	
Evening time spent (4-7 pm; 7-10 pm; 10 pm-12 am)	~5% 4-7, 7-10 ~4.75% 10-12	~5.75% 4-7, 7-10	~1% 4-7, 7-10	N/A (not top 10 predictor)	Evening time spent in the top 10; late night use relatively less important
Passive sessions (overall, and evening 4-7 pm; 7-10 pm)	~5% overall ~4.5% evening	~5.5%	~.6%	~.5-.6%	

Slide 50

19 is there no video chaining behavioral analog for FB? or just didn't rise as a predictor?
Elena Davis, 12/22/2022

20 basically wondering how comparable these are, re: whether things that are MISSING from one or the other platform are meaningful in their absence
Elena Davis, 12/22/2022

18 capturing top 10 predictors
Elena Davis, 12/22/2022

17 is this true? did direct sessions just not come up for IG? or is there no analogous predictor
Elena Davis, 12/23/2022

11 I need to check on it. I think DM's may be capturing a ton of the IG push notification relationship.

It also appears push notifications may be much more prominent on FB (i'm seeing 70% of the surveyed FB sample having a session originating from push).

This isn't a directly comparable thing for a bunch of reasons, but Quiet mode experiment on IG only has like 8% of sessions originating from push notifications.

So how much both of those play in is important
Josh [redacted] 12/23/2022

Global behavior correlations - FB adults

Correlations	Attentiveness	Capability
surface_visits	-0.15	-0.04
engagements	-0.07	-0.02
time_spent	-0.15	-0.03
push_notif_session	-0.15	-0.06
time_to_sess_from_push	0.00	0.00
late_night_ts	-0.09	-0.02
ts_4_7	-0.13	-0.01
ts_7_10	-0.13	-0.03
ts_10_12	-0.11	-0.03
late_night_sessions	-0.13	-0.04
no_eng_sessions	-0.14	-0.04
late_night_high_surface_switching_sessions	-0.10	-0.03
notif_checks	-0.10	-0.02
short_sessions	-0.12	-0.04
repeat_no_eng_session	-0.14	-0.06
late_night_no_eng_sessions	-0.09	-0.03
no_engagement_4_7	-0.13	-0.04
no_engagement_7_10	-0.13	-0.04
no_engagement_10_12	-0.11	-0.04

Global behavior correlations - IG youth

Correlations	Attentiveness	Capability
ts	-0.23	-0.11
ts_4_7	-0.10	-0.06
ts_4_7_passive	-0.07	-0.05
ts_7_10	-0.19	-0.10
ts_7_10_passive	-0.14	-0.09
ts_10_12	-0.18	-0.09
ts_10_12_passive	-0.14	-0.08
ts_late_night	-0.14	-0.07
ts_late_night_passive	-0.11	-0.07
no_eng_repeat_sessions	-0.14	-0.10
short_sessions_count	-0.11	-0.08
late_night_no_engagement_sessions	-0.09	-0.07
sessions	-0.18	-0.11
sessions_exceed_2hr	-0.07	-0.03
avg_max_chaining_depth	-0.04	0.01
sum_chaining_depth_max	-0.07	0.00
post_impressions_chaining	-0.08	-0.01
media_impressions_chaining	-0.09	-0.02

Global behavior correlations - IG adults

Correlations	Attentiveness	Capability
ts	-0.22	-0.11
ts_4_7	-0.10	-0.05
ts_4_7_passive	-0.08	-0.04
ts_7_10	-0.18	-0.10
ts_7_10_passive	-0.15	-0.08
ts_10_12	-0.18	-0.08
ts_10_12_passive	-0.14	-0.07
ts_late_night	-0.13	-0.08
ts_late_night_passive	-0.11	-0.07
no_eng_repeat_sessions	-0.17	-0.10
short_sessions_count	-0.15	-0.09
late_night_no_engagement_sessions	-0.10	-0.08
sessions	-0.22	-0.10
sessions_exceed_2hr	-0.04	-0.02
avg_max_chaining_depth	-0.05	-0.02
sum_chaining_depth_max	-0.04	-0.02
post_impressions_chaining	-0.06	-0.04
media_impressions_chaining	-0.07	-0.05

Overall behavior patterns - Instagram

21 havioral signal	**Low attention** Strength of associations		**Low capability** Strength of associations		
	Teens	Adults	Teens	Adults	Overall predictor strength for linking on-platform signals with low control are 2-3x stronger for predicting low attention than low capability Predictors are largely similar across adults and teens
Total time spent	~9%	~9%	~3.5%	~3.5%	Time spent and sessions (as indices of higher overall use) consistently a top predictor
Sessions (number)	~8%	~9%	~3.5%	~3%	
Evening time spent (7-10 pm; 10 pm-12 am)	~7% each	~7% each	~3% each	~3% 7-10 ~2.5% 10-12	Evening time spent, both overall and passive time, as a stronger predictor than late night use (particularly stronger than passive late night use)
Passive evening time spent (7-10 pm; 10 pm-12 am)	~6% each	~6% each	~3% each	~2.75% 7-10 ~2% 10-12	
Late night time spent (12 am - 4 am) (overall and passive)	~6% overall ~4.5% passive	~5.5% overall ~4.75% passive	~2.5% each	~2.5% each	
No engagement	~5.5%	~7%	~3%	~3%	No engagement repeat sessions relatively stronger as a

Slide 54

21 one thought as comparing the strength of associations across these two clusters... any comparability to Carlos's original work? are we doing better b
we have articulated a separate attention construct?
Elena Davis, 12/23/2022

23

Next steps

Slide 55

22 @[redacted]meta.com I think this is what we need to clearly articulate for next week
Elena Davis, 2/4/2023

23 describe what measurement next steps would be -- recommendations for targeting, goaling, heuristics
and 'here's how we propose we engage with you for next steps' - defining product improvements, clarifying goals, expanding solutions
Elena Davis, 2/4/2023

Measurement next steps to get to updated heuristics/metric goals

These global behaviors can be used for targeting/triggering thresholds (?already in Feed Pause)

Next steps involve iterating on different versions of combined signals + thresholds to optimize precision/recall of perceived low control:

?–also include team priorities - eg., late night use

1) [Here are the heuristic identification] Stacking of individual behaviors –
   a) Alternative decision tree models that are abbreviated: simplify the metrics included (e.g., between 3-5 predictors); ensure that each are independently meaningful
   b) Can use those decision tree branching to determine thresholds
   c) This will give a smaller set of behaviors that are all together the best predictors

1) Then redo thresholding based on what specific control product would be; e.g., if we were *only* targeting video chaining, what would the right threshold be? (e.g., behavior thresholds conditional on other targeting factors)
   a) This would inform product strategy for the contexts in which features are offered

1) Opportunity size the potential impact of improved targeting for QM and TaB would be, on behavior change outcomes (even beyond only late night use); e.g., given a certain targeting threshold, what % of TAM is reached? e.g., refine what a reasonable behavior change goal would be (...leveraging efficacy of products to date;) –
   a) E.g., extending Quiet mode to earlier hours
   b) E.g., Take a Break triggering

Slide 56

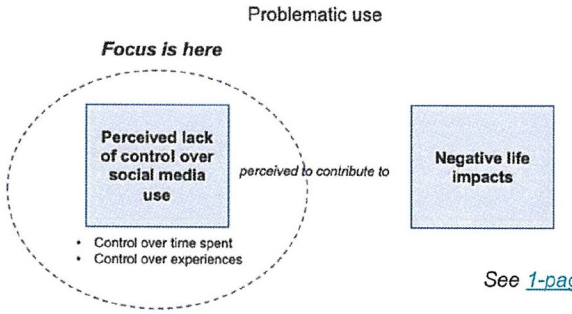
24 updated short list of independent heuristics
Elena Davis, 2/7/2023

25 new opportunities - product recommendations for expansion in product to other surfaces/triggering
Elena Davis, 2/7/2023

26 for new investment
Elena Davis, 2/7/2023

Focusing on control increases our measurement precision

This supports our approach to focus on control: from both a conceptual and measurement standpoint, perceived control over platform use is more proximal to the actual on-platform experience, including behaviors that we can measure and potentially change.



See 1-pager here

Alternative model

	Baseline	Precision	Recall	AUC
Former PU model on FB Outcome: low control + negative life impact	~0.11	0.17 all ages	0.73 all ages	0.66 all ages
Current PU model on FB Outcome: low attentiveness	~0.3-0.35	0.67 youth 0.57 adults	0.37 youth 0.44 adults	0.74 youth 0.69 adults
Current PU model on FB Outcome: low capability	~0.15-0.2	0.64 youth 0.46 adults	0.25 youth 0.24 adults	0.76 youth 0.73 adults

Background

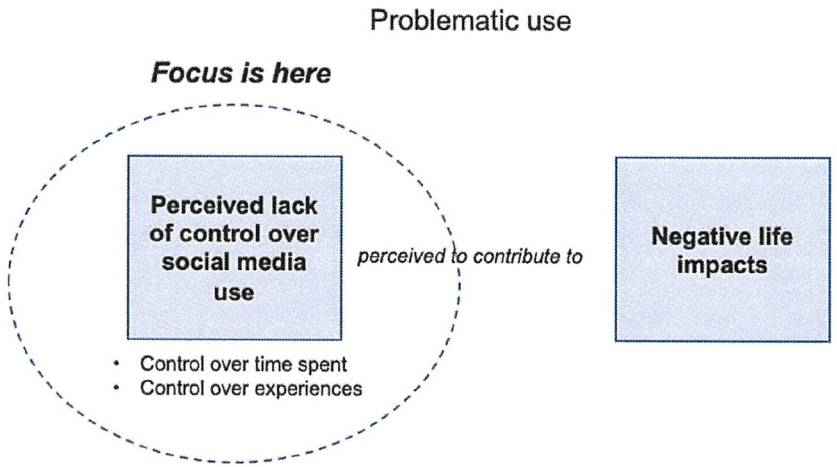
Problematic Use Definition & Success State

Problematic use is:

1. a perceived lack of control over usage of a technology or social media platform

and

1. this lack of control is perceived to lead to negative life impacts (e.g., in domains such as relationships, work/school, sleep, etc)



Overall success:

People have the tools that they need and are empowered to control their time and experiences on our platforms.

See 1-pager here

What does potential problematic use look like?

Perceived (lack of) control over social media use[1]:

- **Control over time spent**: includes overall time spent, when time is spent, how often they spend time or check the platform
- **Control over experiences**: includes content consumption patterns and interactions with others on our platforms

Key control constructs necessary for behavior change: Awareness of behavior, Attentiveness to behavior, Intentionality and choice (capability to change)

[1]See 1-pager here

Survey details

Statistical support for composite scores

	Facebook	Instagram
Do the items follow a pattern of 2 clusters of 'control'? Confirmatory factor analysis *Note: thresholds for strong fit are CFI > 0.95 and RMSEA < .05*	Yes 2 factor model has strong fit for both cohorts: • Adult (CFI = 0.995; RMSEA = 0.037) • Youth (CFI = 0.997; RMSEA = 0.026)	Yes 2 factor model has strong fit for both cohorts: • Adult (CFI = 0.995; RMSEA = 0.039) • Youth (CFI = 0.992; RMSEA = 0.042)
Is the pattern of 2 clusters the same across cohorts? *Invariance testing* *Note: Non-significant Chisq difference means models are statistically the same*	Yes - partially Same clusters are present for both cohorts (adults + youth) and both platforms, but the contributions of specific items to the clusters are different • Configural invariance (same clusters): CFI = 0.991, RMSEA = 0.038 • Metric invariance (same loadings of each items within cluster): chisq difference (12) = 150.75 < .001	
Can we form composite scores for these 2 clusters? *Internal reliability Cronbach's alpha* *Note: threshold for internally reliable is alpha >0.7*	Yes Attentiveness composite • Adult alpha = 0.77 • Youth alpha = 0.75 Capability composite • Adult alpha = 0.79 • Youth alpha = 0.77	Yes Attentiveness composite • Adult alpha = 0.79 • Youth alpha = 0.78 Capability composite • Adult alpha = 0.74 • Youth alpha = 0.74

Results details: descriptives

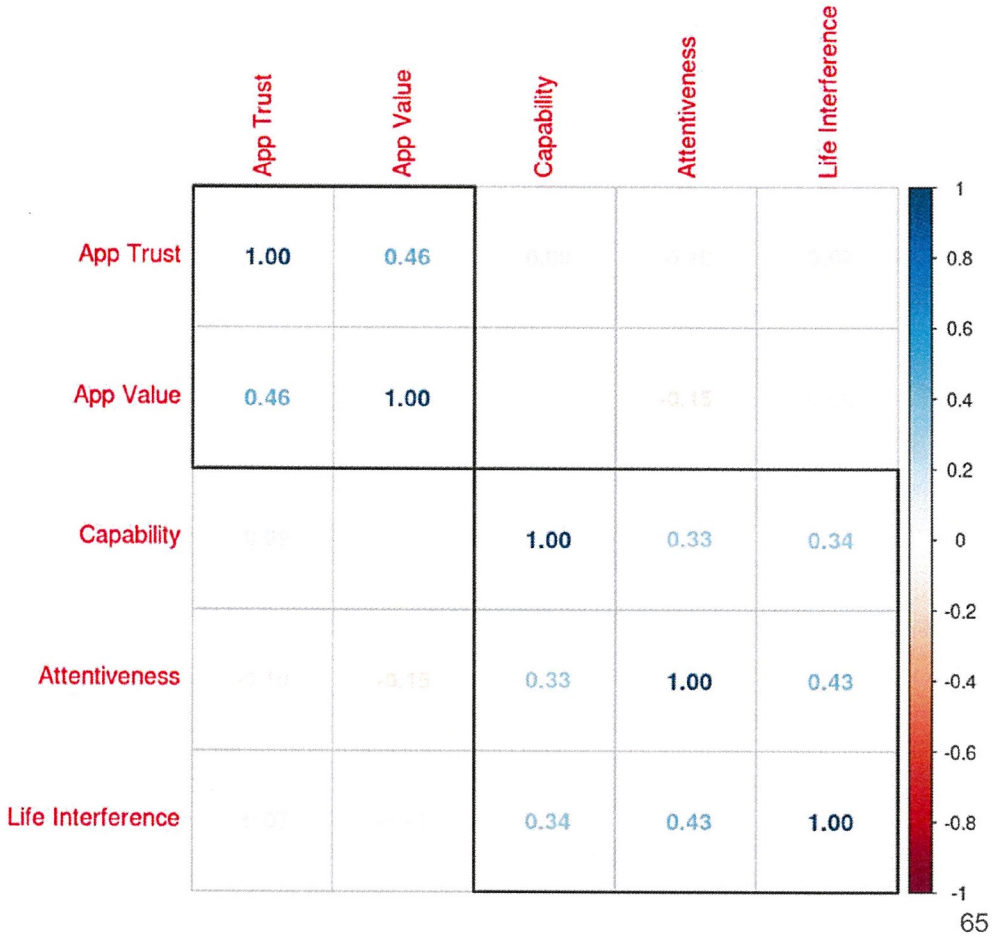
Descriptives - Facebook

Higher scores = 'better' (more control or less life interference). Results weighted for age, gender, country (within sample set), active use. See detail.

	Average 'attentiveness' score *(composite of 3 items)*	% reporting lowest attentiveness *(lowest control response to 1 or more attentiveness items)*	% reporting lowest attentiveness *(<=2 on the composite)*	Association with life interference *(lowest 2 responses of life interference "all the time' + 'often')*	Average 'capability' score *(composite of 3 items)*	% reporting lowest capability *(lowest control response to 1 or more capability items)*	% reporting lowest capability *(<=2 on composite)*	Association with life interference *(reporting <=2 on life interference item)*	% reporting high life interference *(lowest 2 responses of life interference "all the time' + 'often')*	Overlap between low attentiveness & low capability *(proportion reporting both)*
Adults	2.68 (1.05) [95% CI: 2.67-2.69]	32% [95% CI: 31.8-32.8%]	34.5% [95% CI: 34.1-35.0%]	As attention decreases, 2.7x as likely to report life interference Adults with low attention report 23% higher life interference (4.2 vs. 3.3)	3.46 (1.11) [95% CI: 3.45-3.47]	17% [95% CI: 16.1-16.9%]	15.6% [95% CI: 15.2-16.0%]	As capability decreases, 1.8x as likely to report life interference Adults with low capability report 23% higher life interference (4.0 vs. 3.1)	15.4% [95% CI: 15.1-15.9%]	7.3% [95% CI: 7.1-7.6%]
Youth	2.87 (1.16) [95% CI: 2.86-2.88]	34% [95% CI: 33.9-34.8%]	30.9% [95% CI: 30.5-31.4%]	As attention decreases, 2.4x as likely to report life interference Youth with low attention report 25% higher life interference (3.9 vs. 2.9)	3.41 (1.19) [95% CI: 3.39-3.42]	20% [95% CI: 19.9-20.1%]	19.1% [95% CI: 18.6-19.5%]	As capability decreases, 1.7x as likely to report life interference Youth with low capability report 23% higher life interference (3.8 vs. 2.9)	21.4% [95% CI: 20.9-22.0%]	7.8% [95% CI: 7.5-8.0%]
Comparison	Controlling for country, adults more likely to report lower attentiveness (p<.001 in regression)	Chisq(1)=22.2 p<.001 Youth more likely to report *lowest* attentiveness	Chisq(1)=70.3 p<.001 Adults more likely to report overall low attentiveness		Controlling for country, youth more likely to report lower capability (p<.001 in regression)	Chisq(1)=97,9 p<.001 Youth more likely to report *lowest* capability	Chisq(1)=85.8 p<.001 Youth more likely to report overall low capability		Chisq(1)= 193.7, p<.001 Youth more likely to report high life interference	Chisq(1)= 3.26, p=.07 (not significant)

Facebook Youth Correlations: Composites + Key variables

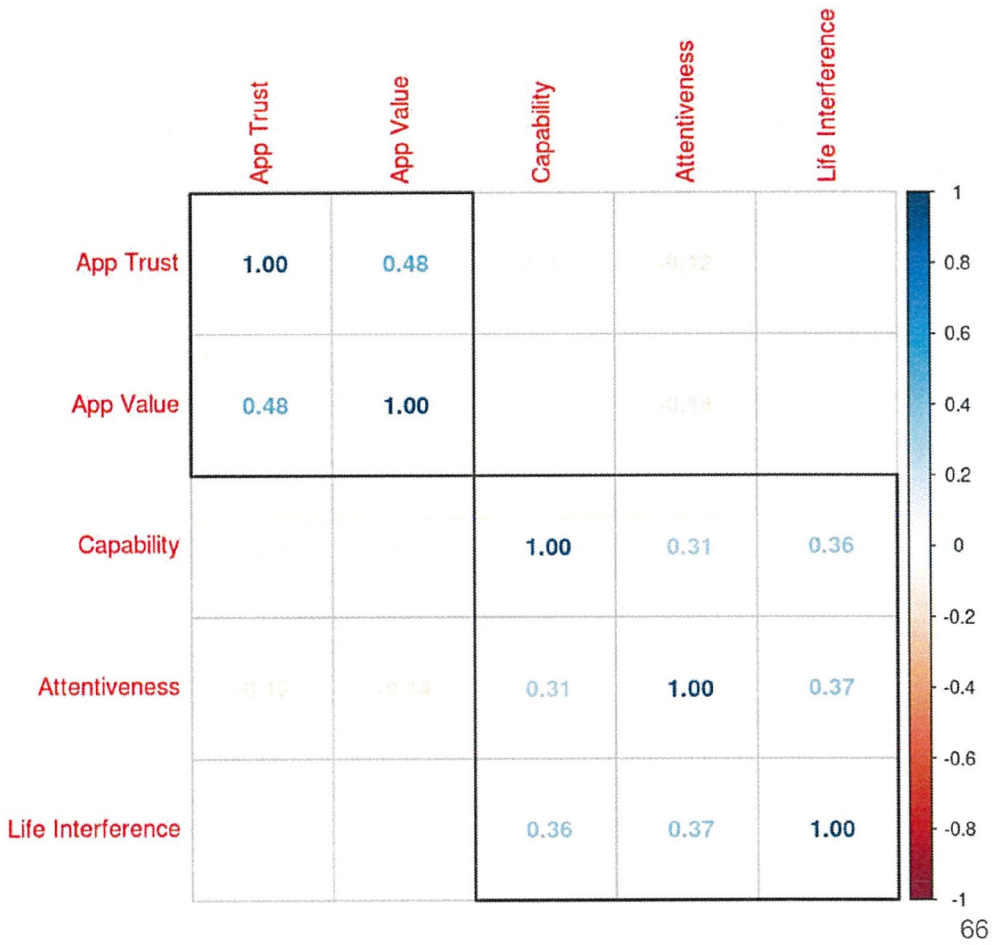
- Composites moderately correlated with each other
- Both moderately correlated with life interference; **attentiveness slightly more so than capability**
- Fairly minimal association with value/trust
- Value and trust moderately correlated



65

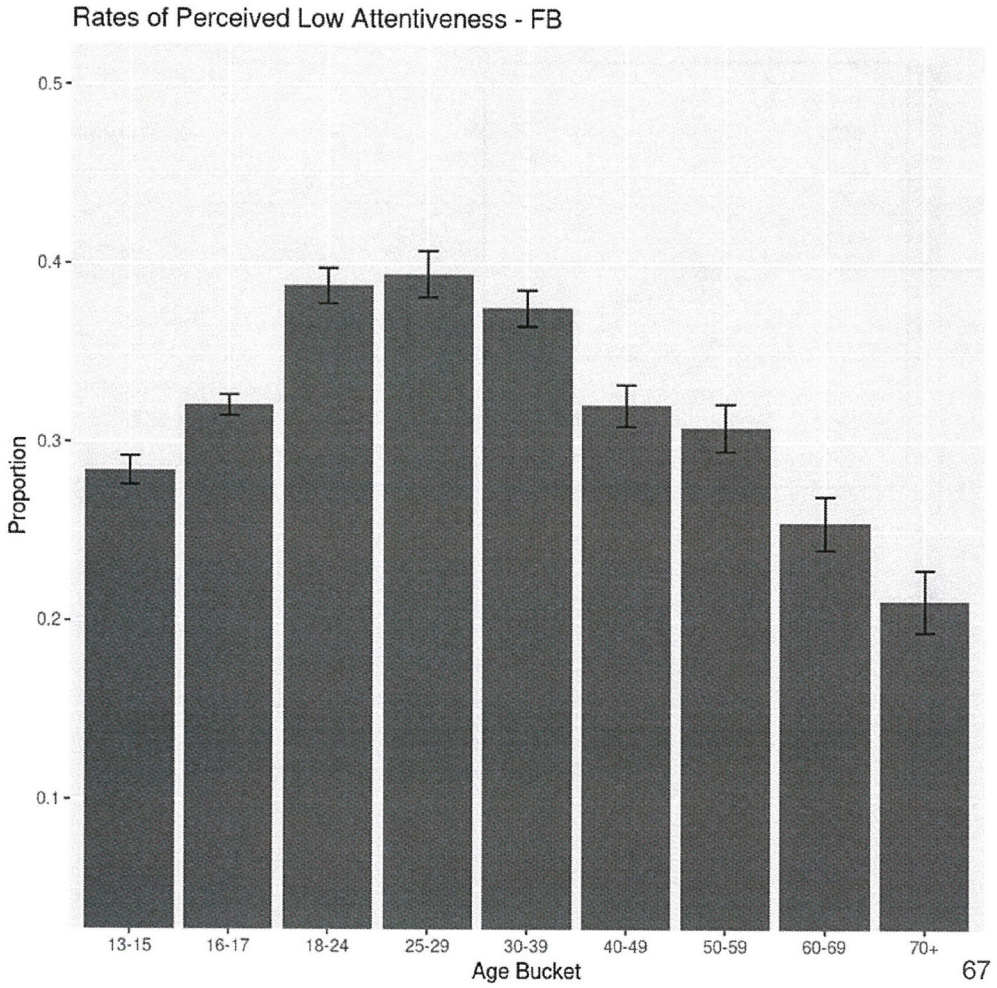
Facebook Adult Correlations: Composites + Key variables

- Composites moderately correlated with each other
- Both moderately correlated with life interference; **attentiveness same as capability**
- Fairly minimal association with value/trust
- Value and trust moderately correlated

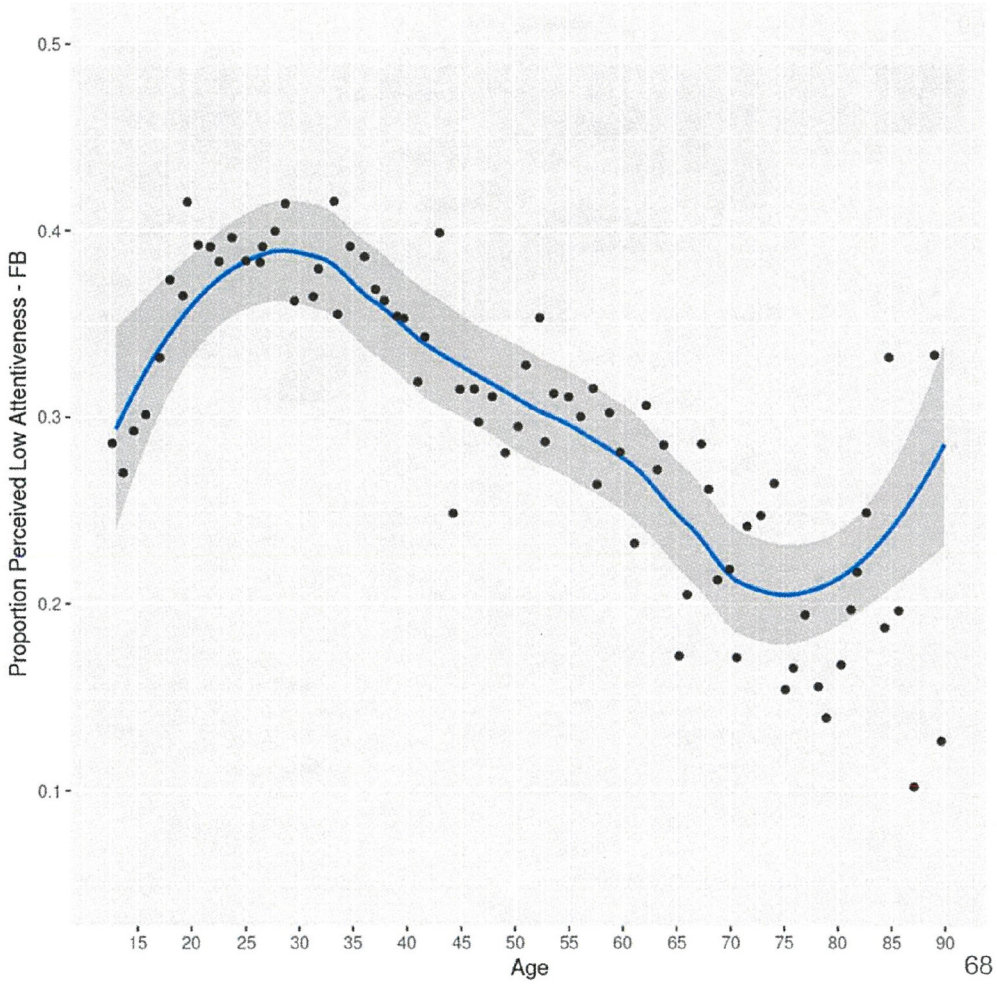


66

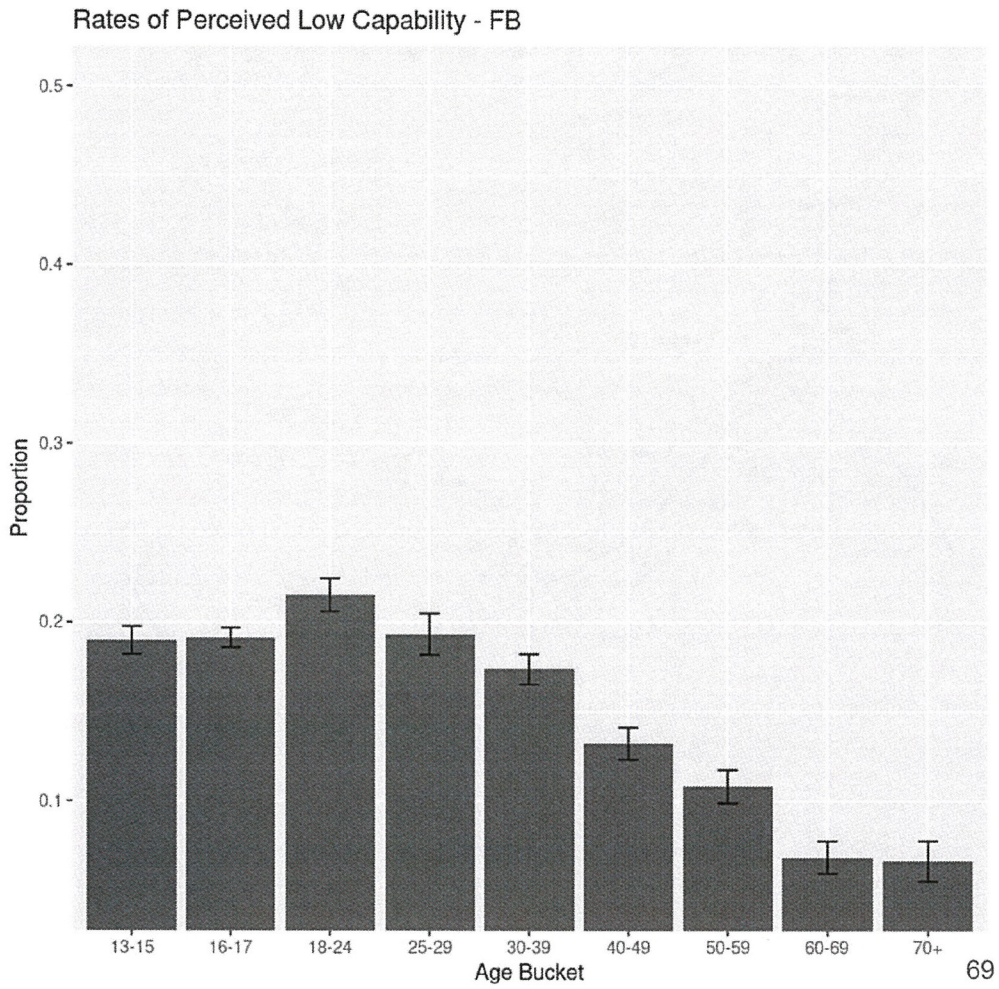
Facebook: Perceived low attentiveness by age



Facebook: Perceived low attentiveness by age

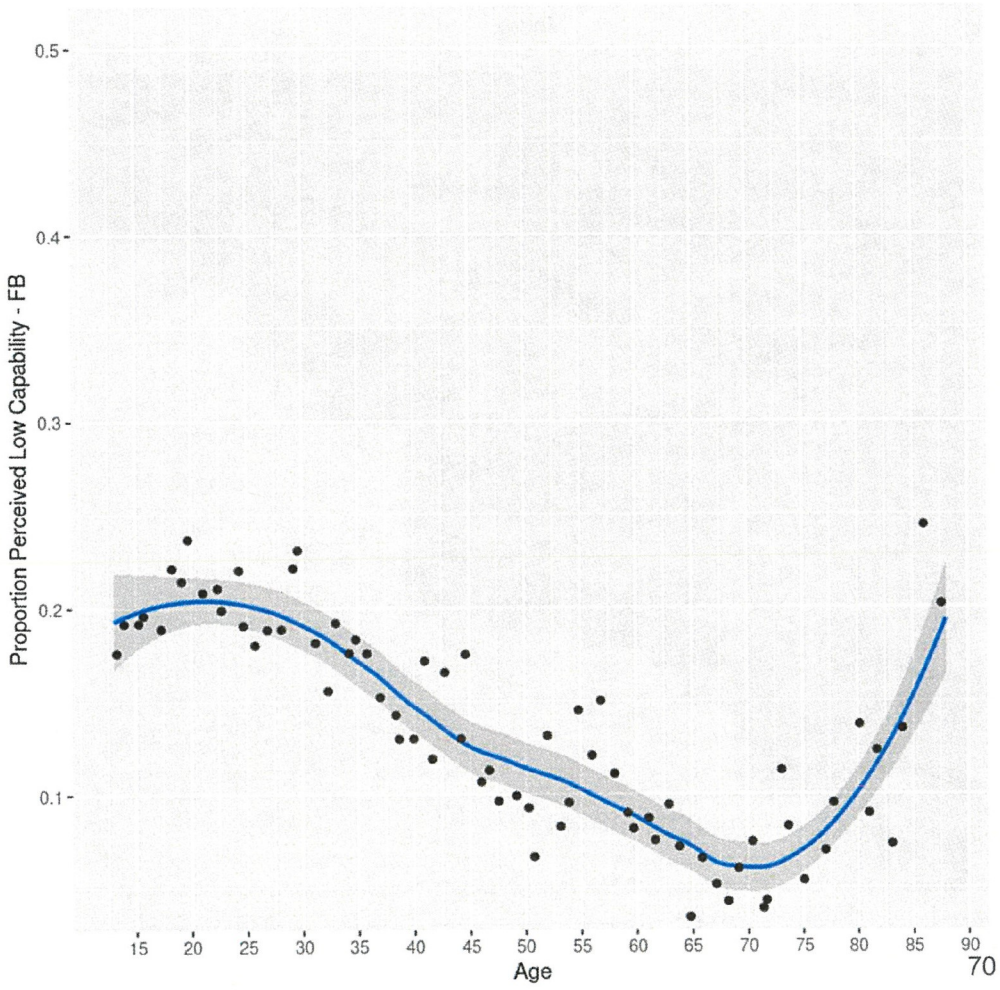


Facebook: Perceived low capability by age

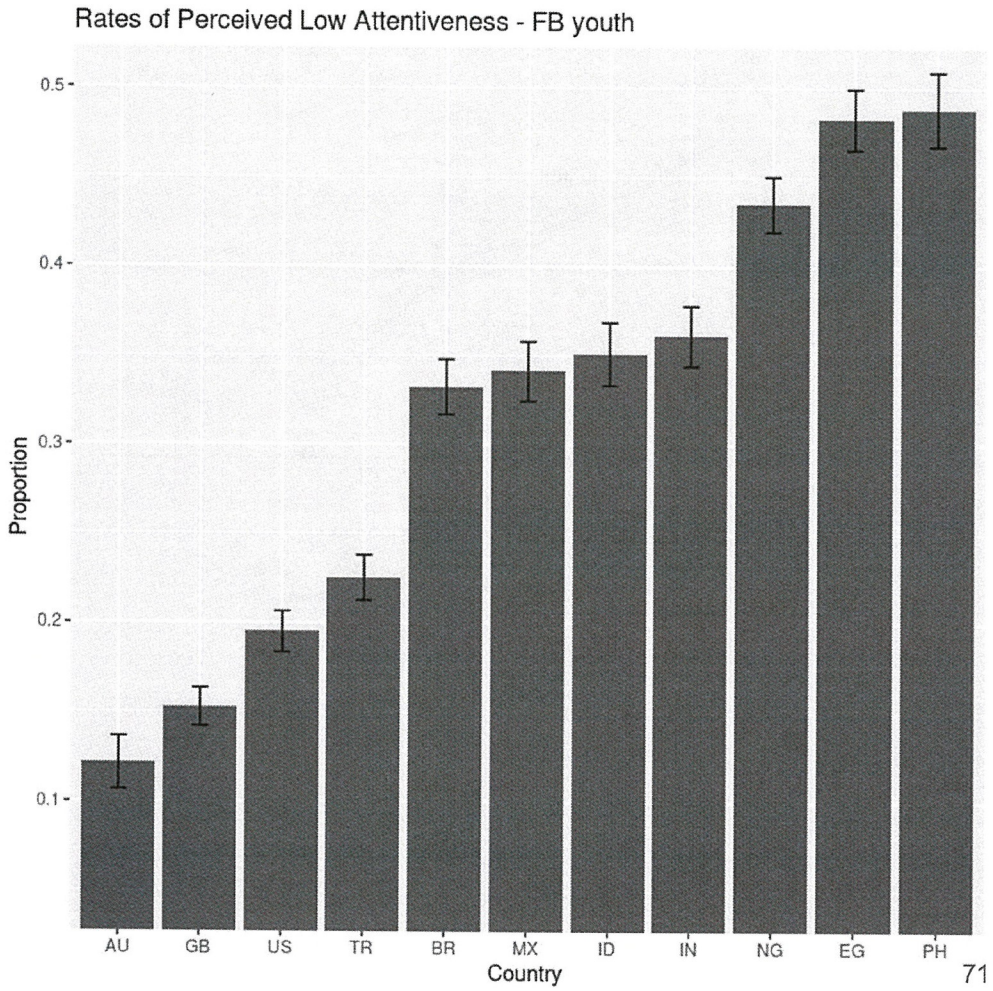


69

Facebook: Perceived low capability by age

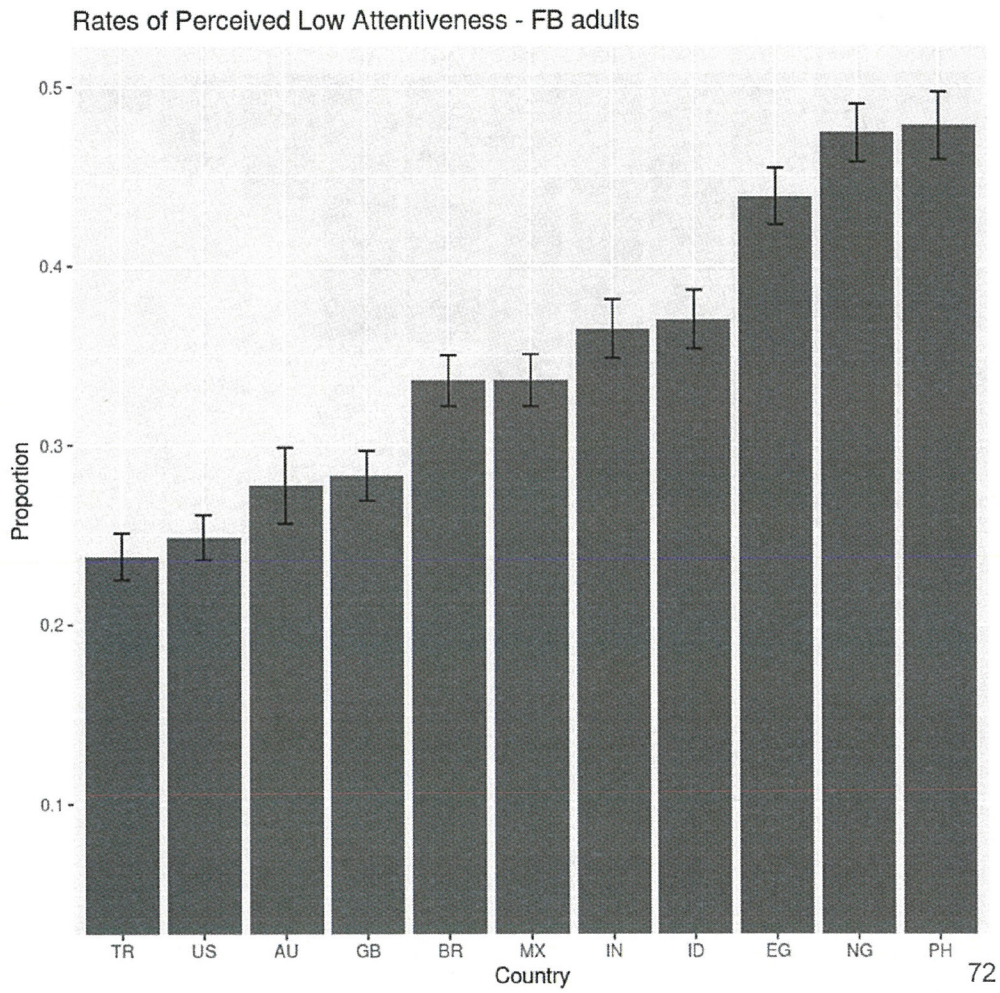


Facebook Youth: Perceived low attentiveness by country

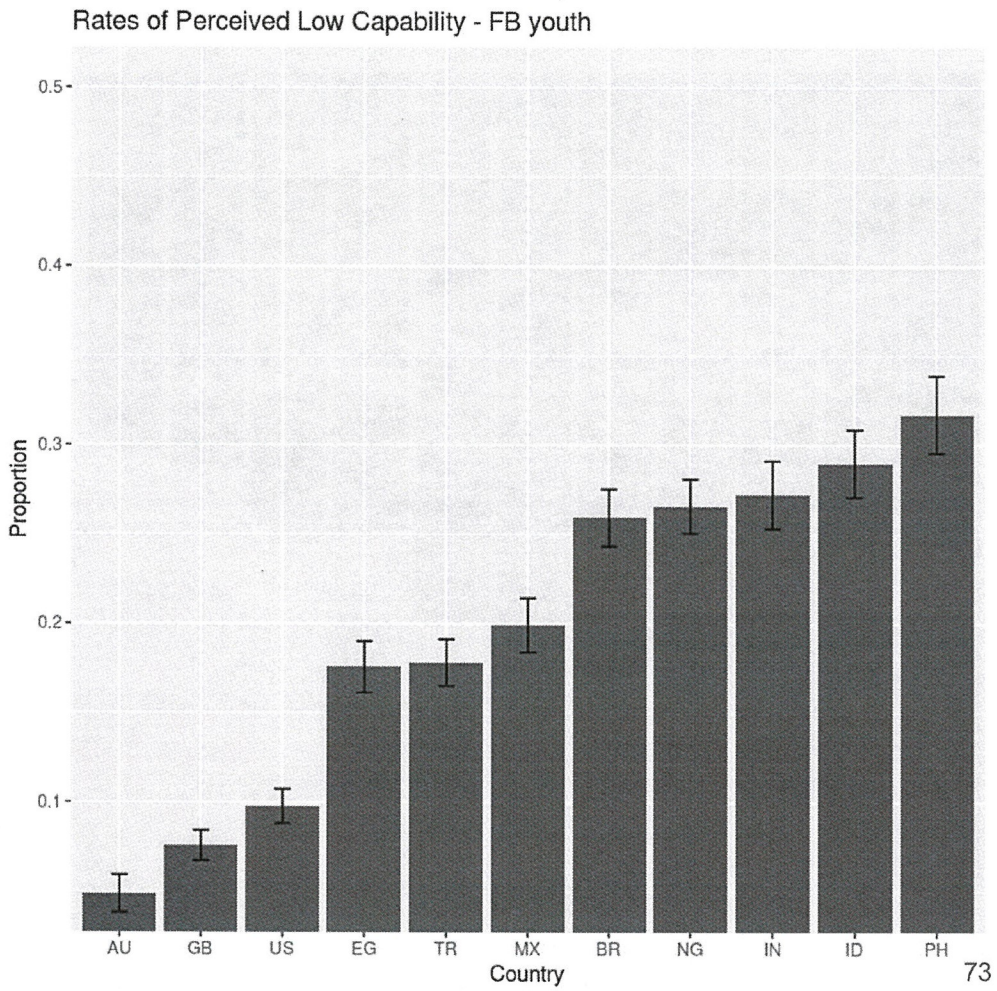


71

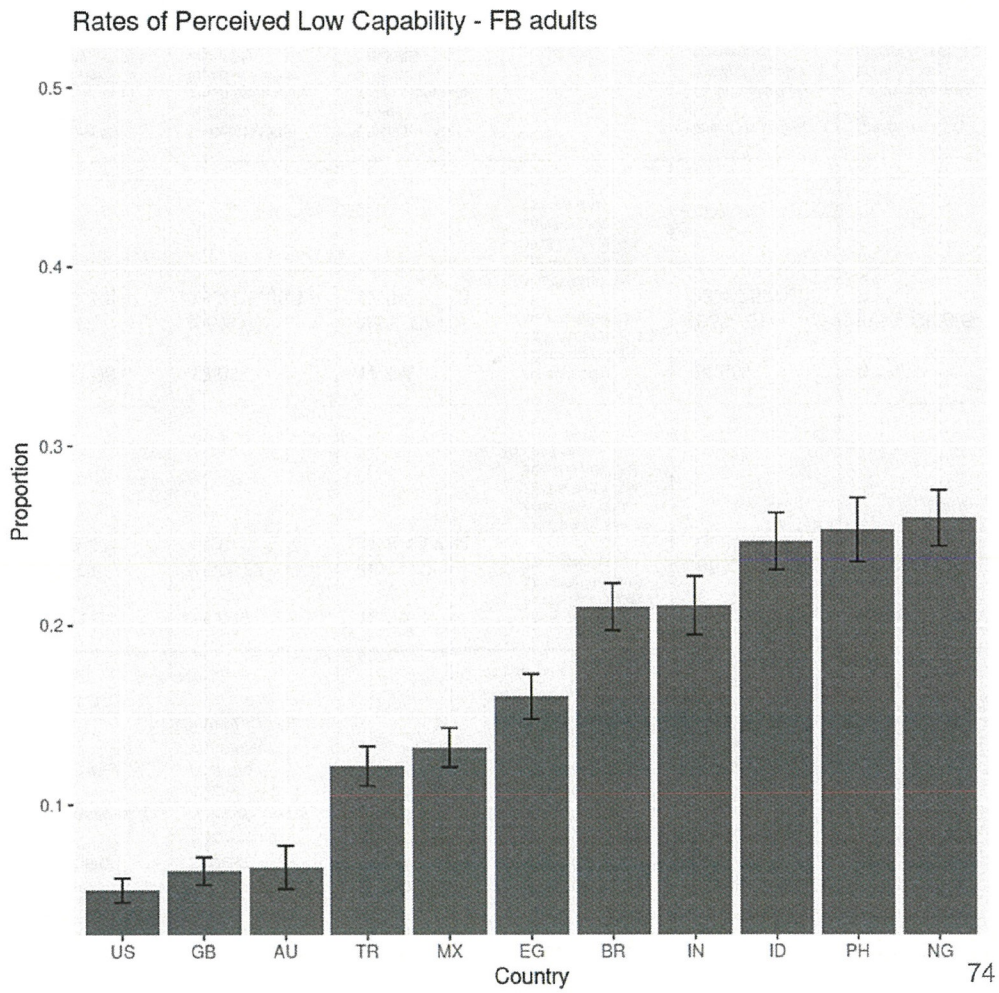
Facebook Adults: Perceived low attentiveness by country



Facebook Youth: Perceived low capability by country



Facebook Adults: Perceived low capability by country



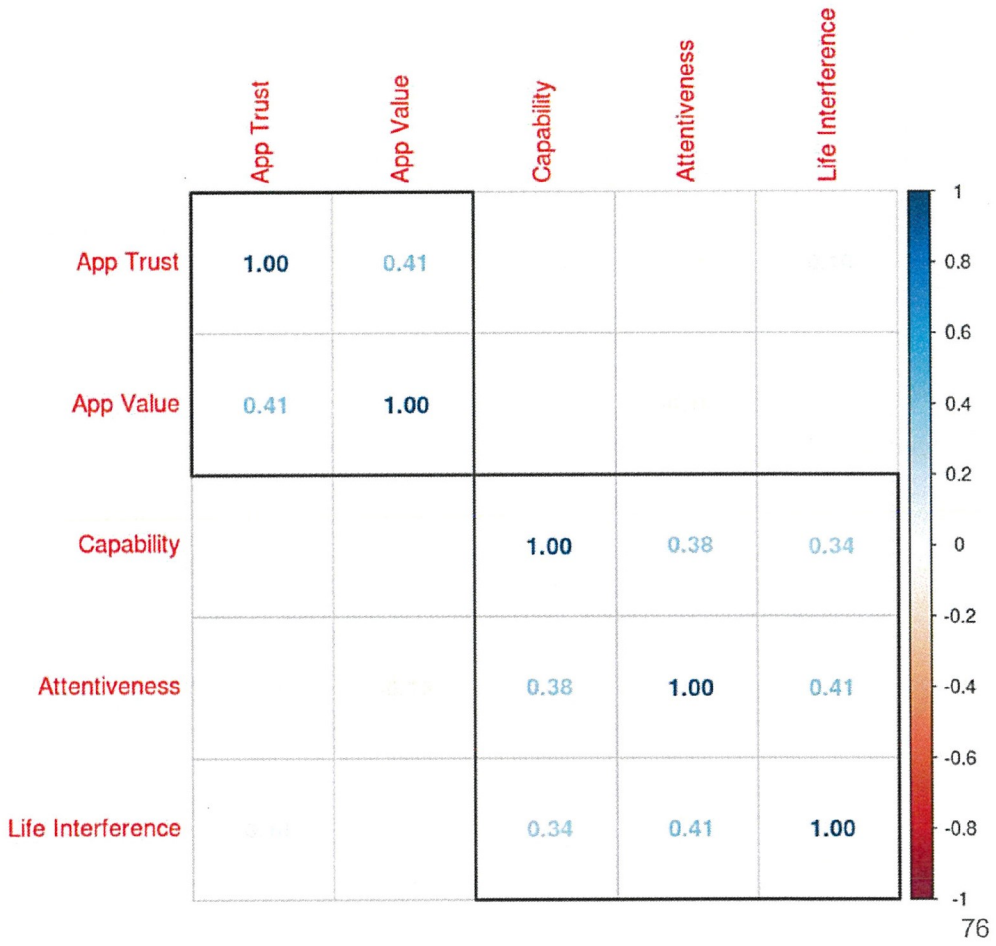
Descriptives - Instagram

Higher scores = 'better' (more control or less life interference). Results weighted for age, gender, country (within sample set), active use. See detail.

	Average 'attention' score *(composite of 3 items)*	% reporting lowest attentiveness *(lowest control response to 1 or more attention items)*	% reporting lowest attentiveness *(<=2 on the composite)*	Association with life interference *(lowest 2 responses of life interference "all the time' + 'often')*	Average 'capability' score *(composite of 3 items)*	% reporting lowest capability *(lowest control response to 1 or more capability items)*	% reporting lowest capability *(<=2 on composite)*	Association with life interference *(top 2 responses of life interference "all the time' + 'often')*	% reporting high life interference *(lowest 2 responses of life interference "all the time' + 'often')*	Overlap between low attentiveness & low capability *(proportion reporting both)*
Adults	2.52 (.99) [95% CI: 2.50-2.53]	34.4% [95% CI: 33.9-35.0%]	40.0% [95% CI: 39.7-40.9%]	As attention decreases, 3.0x as likely to report life interference Adults with low attention report 20% higher life interference (4.1 vs. 3.3)	3.46 (1.05) [95% CI: 3.45-3.47]	14.4% [95% CI: 14.0-14.9%]	13.2% [95% CI: 12.8-13.7%]	As capability decreases, 2.0x as likely to report life interference Adults with low capability report 22.5% higher life interference (3.9 vs. 3.0)	16.0% [95% CI: 15.5-16.5%]	8.1% [95% CI: 7.8-8.4%]
Youth	2.59 (1.04) [95% CI: 2.57-2.61]	39.1% [95% CI: 38.2-40.1%]	37.7% [95% CI: 36.7-38.6%]	As attention decreases, 2.7x as likely to report life interference Youth with low attention report 25% higher life interference (3.8 vs. 2.9)	3.27 (1.08) [95% CI: 3.25-3.30]	19.0% [95% CI: 18.2-19.9%]	17.9% [95% CI: 17.0-18.7%]	As capability decreases, 1.9x as likely to report life interference Youth with low capability report 17.5% higher life interference (3.5 vs. 2.8)	24.5% [95% CI: 23.4-25.6%]	9.1% [95% CI: 8.5-9.7%]
Comparison	Controlling for country, adults more likely to report lower attentiveness (p<.001 in regression)	Chisq(1)=98.0 p<.0001 Youth more likely to report *lowest* attentiveness	Chisq(1) = 29.4; p<.001 Adults more likely to report low attentiveness		Controlling for country, youth more likely to report lower capability (p<.001 in regression)	Chisq(1)=133.6 p<.001 Youth more likely to report *lowest* capability	Chisq(1)=145.1, p<.001 Youth more likely to report low capability		Chisq(1)= 328.4, p<.001 Youth more likely to report high life interference	Chisq(1)= 13.7, p<.001 Youth more likely to report low control across both constructs

Instagram Youth Correlations: Composites + Key variables

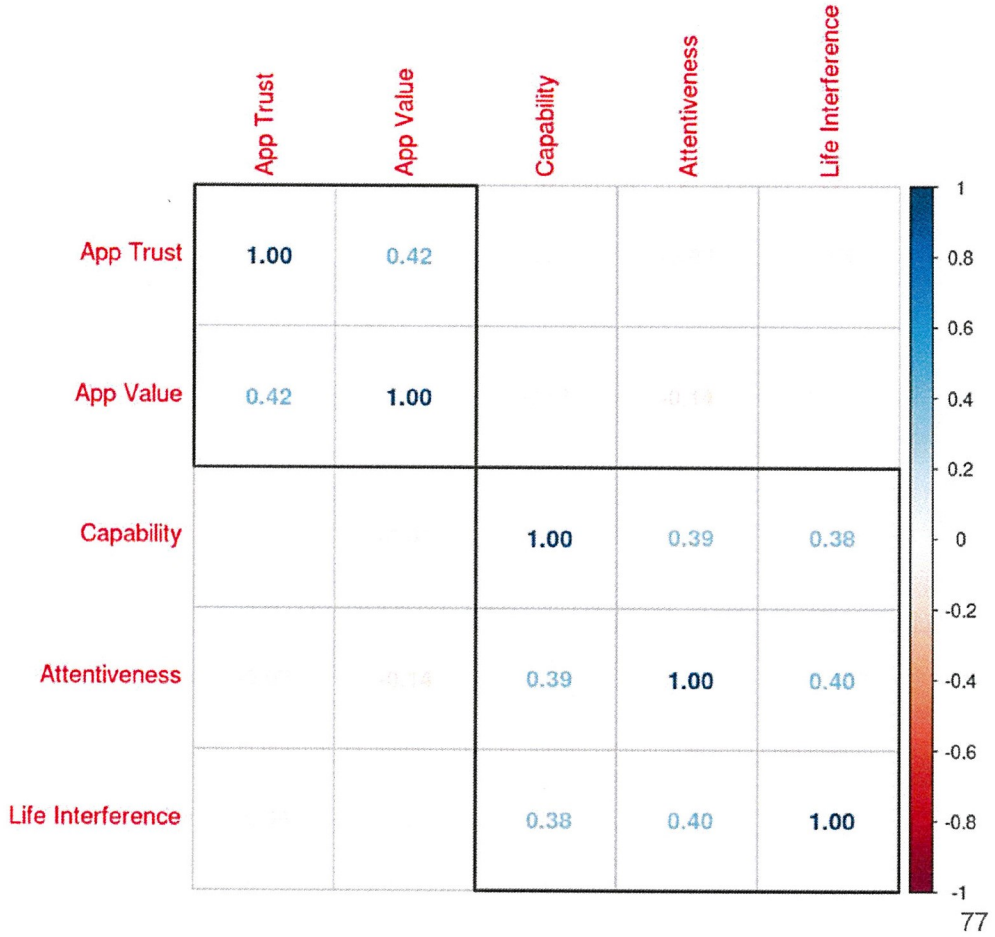
- Composites moderately correlated with each other
- Both moderately correlated with life interference; **attentiveness slightly more so than capability**
- Fairly minimal association with value/trust
- Value and trust moderately correlated



76

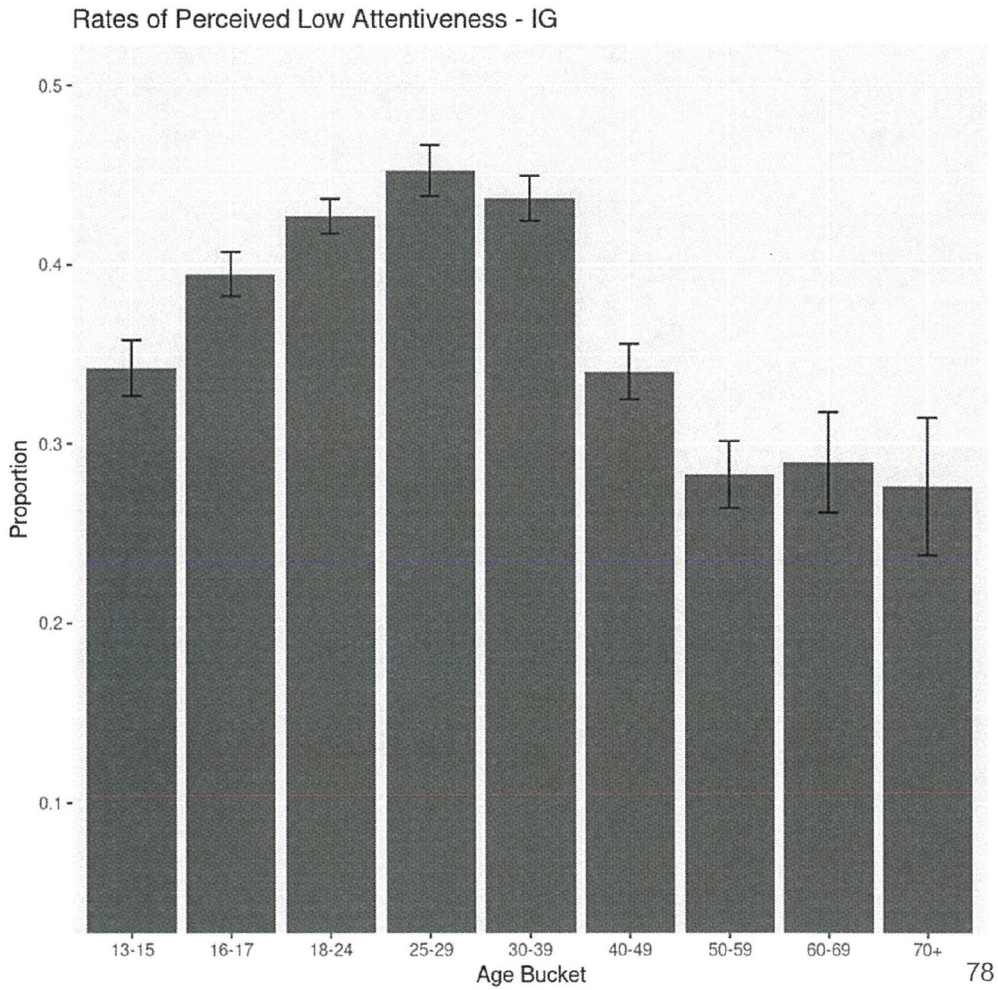
Instagram Adult Correlations: Composites + Key variables

- Composites moderately correlated with each other
- Both moderately correlated with life interference; similar associations with each
- Fairly minimal association with value/trust
- Value and trust moderately correlated

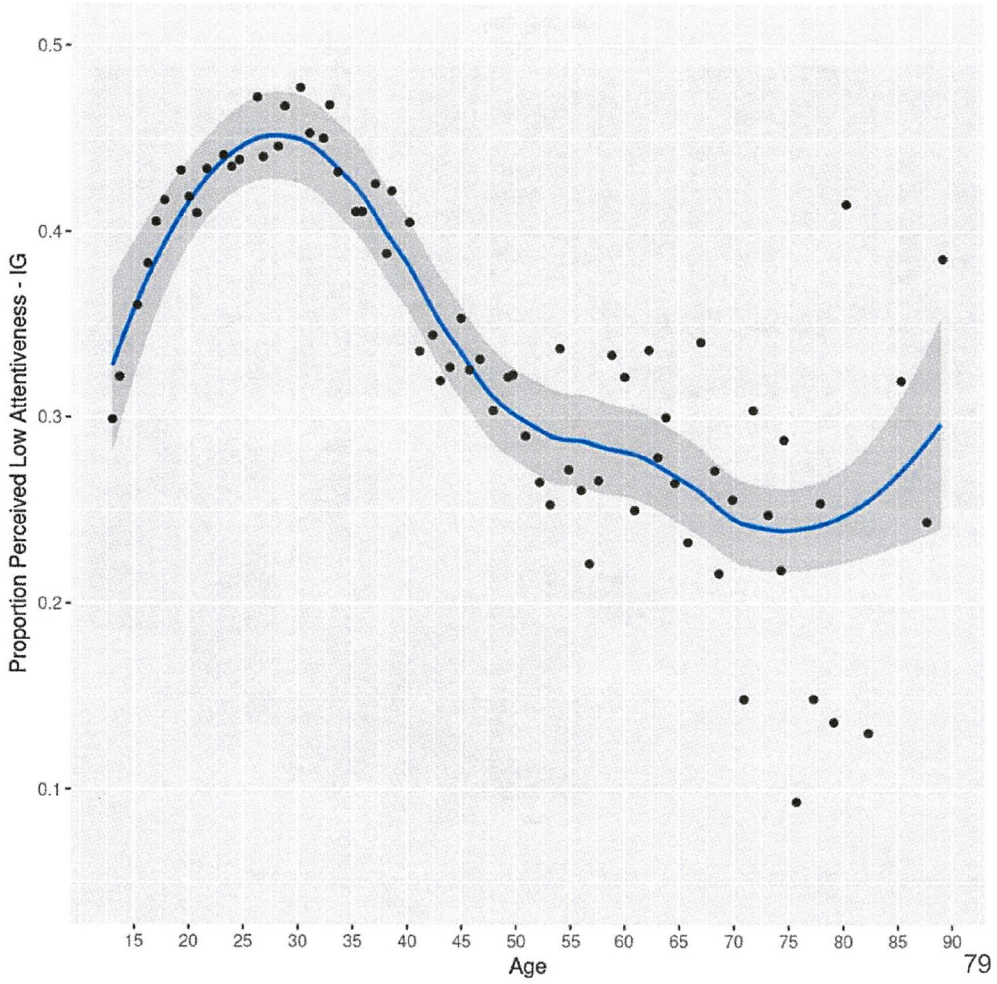


77

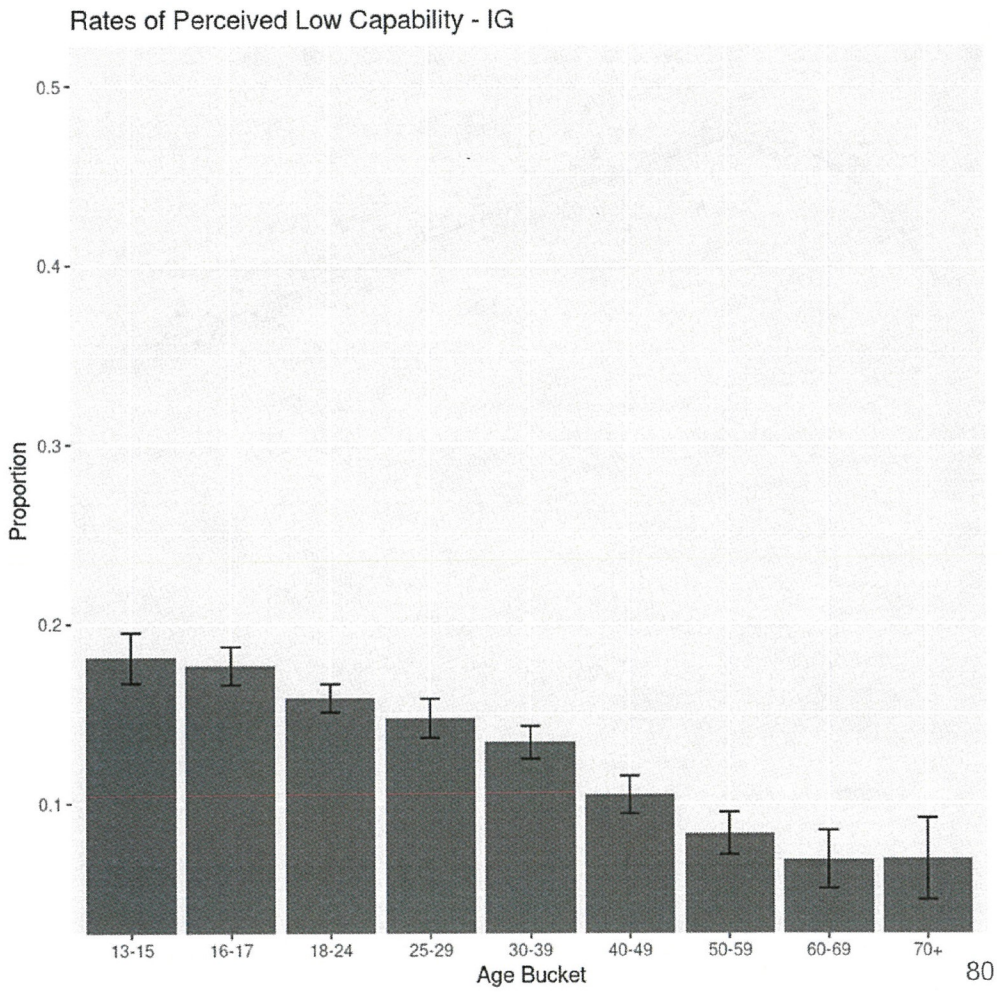
Instagram: Perceived low attentiveness by age



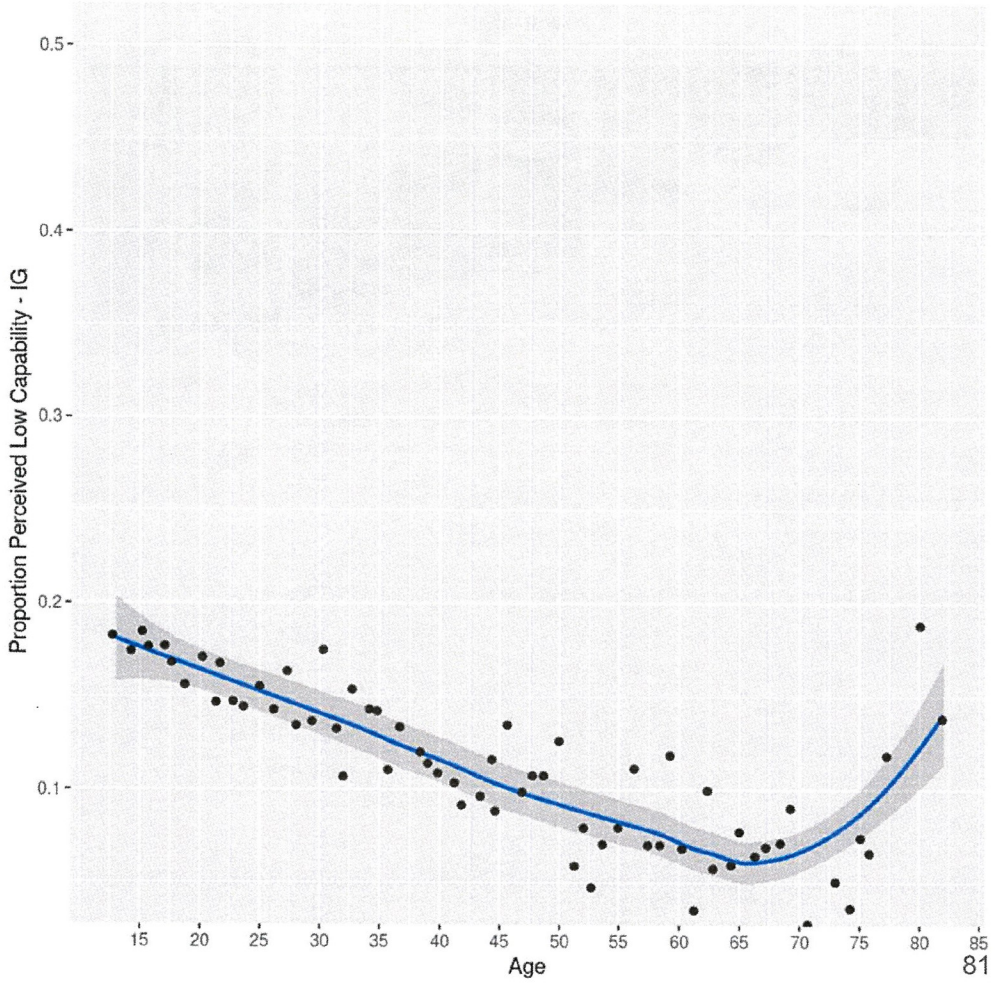
Instagram: Perceived low attentiveness by age



Instagram: Perceived low capability by age

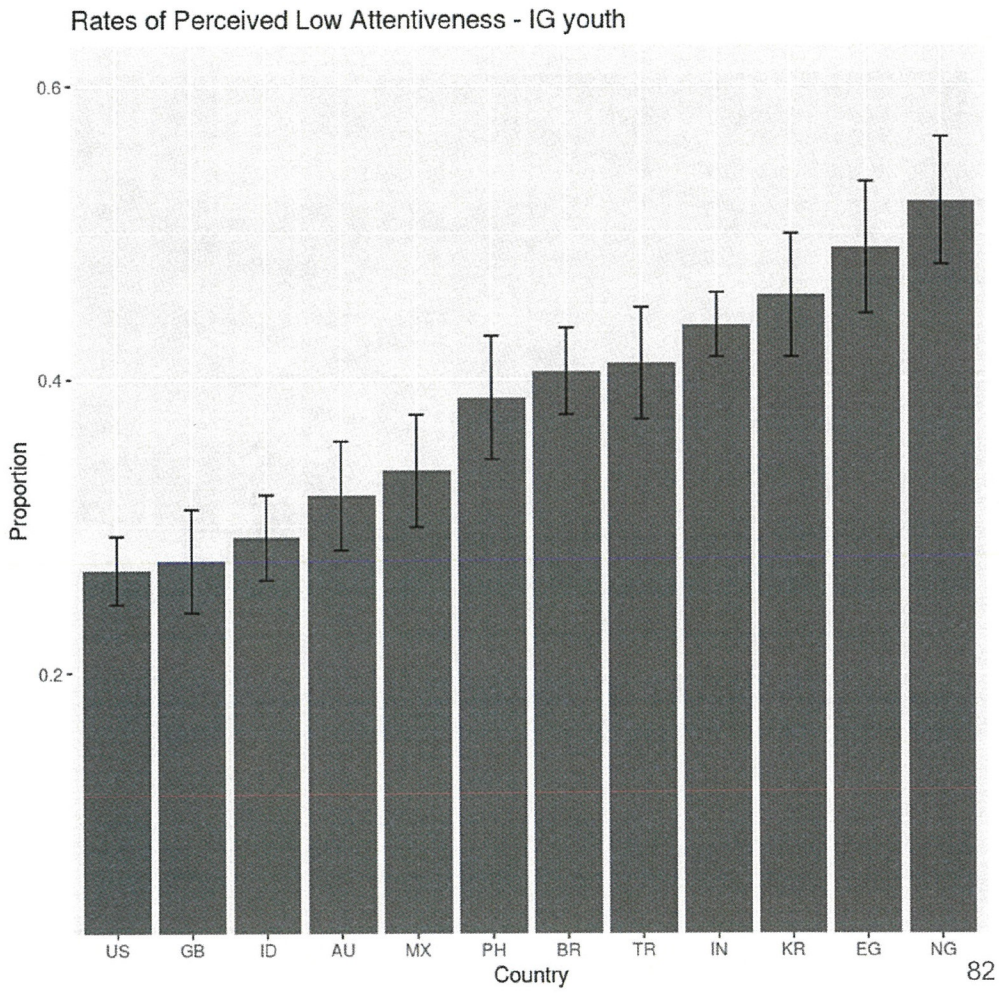


Instagram: Perceived low capability by age

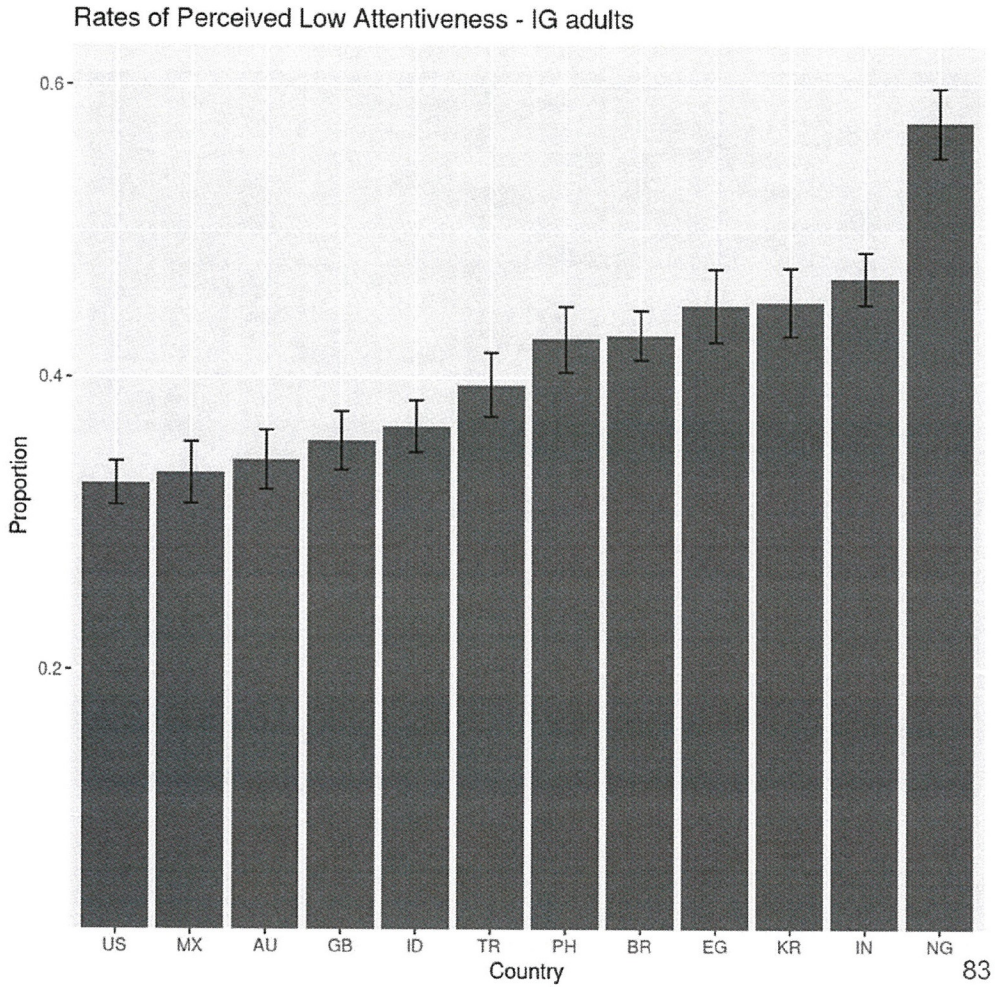


81

Instagram Youth: Perceived low attentiveness by country

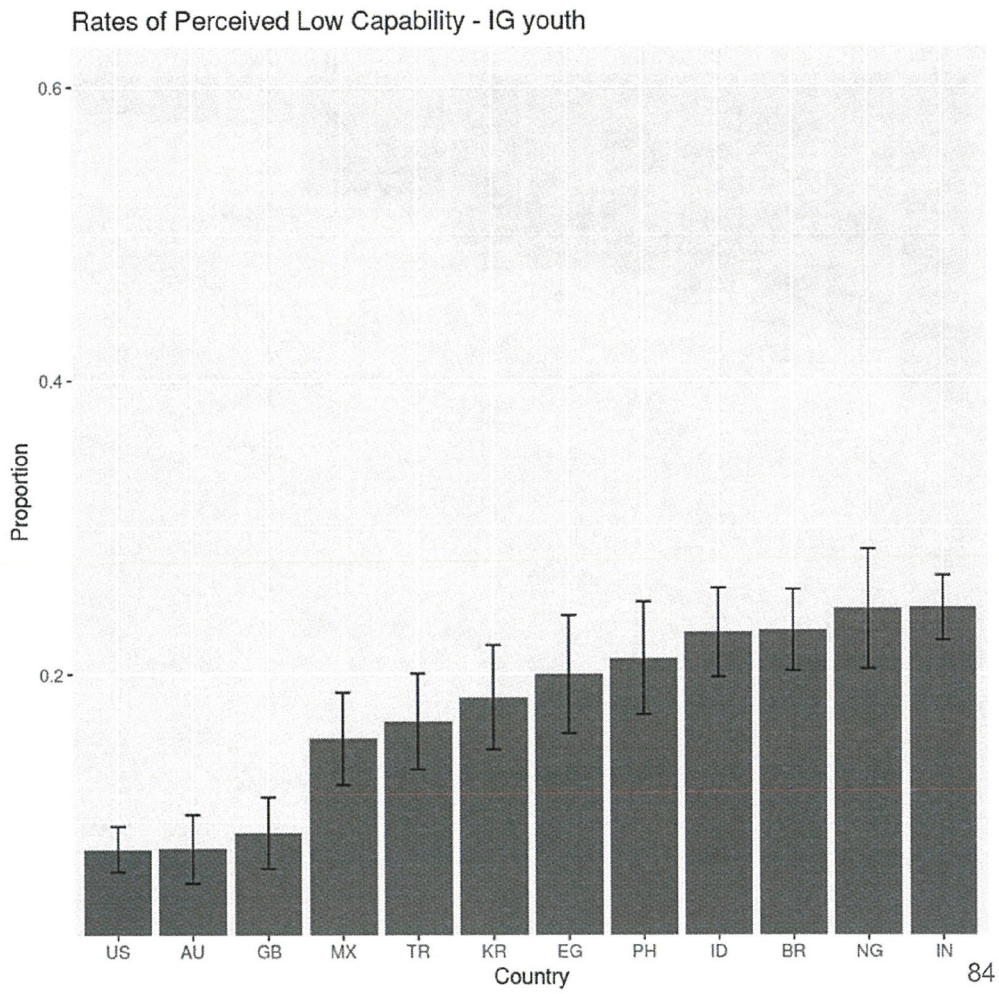


Instagram Adults: Perceived low attentiveness by country

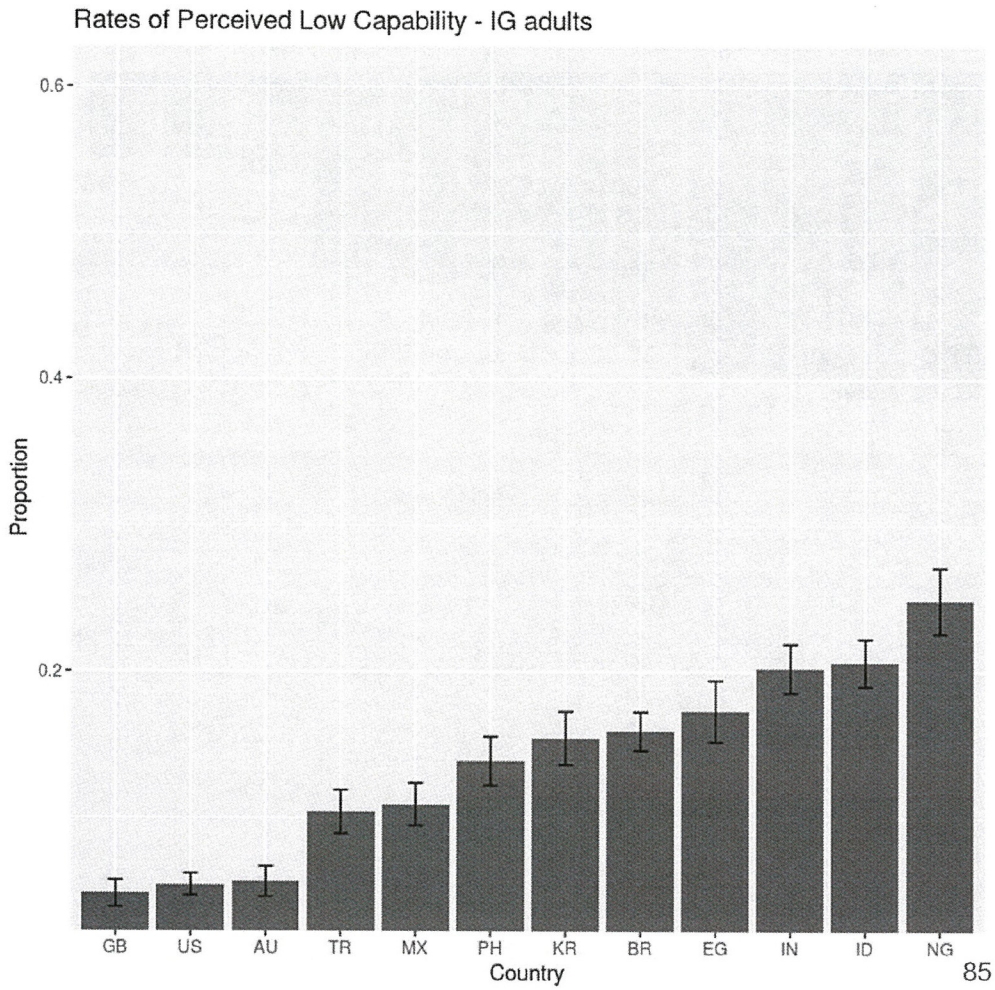


83

Instagram Youth: Perceived low capability by country



Instagram Adults: Perceived low capability by country



Results details: behaviors

Facebook usage (non-causal) associations with perceived control

Product	**Attentiveness** *More time spent within these products associated* with:* **(+) more control**: lower frequency of losing track of time and/or habitual use **(-) less control**: higher frequency of losing track of time and/or habitual use		**Capability** *More time spent within these products associated* with:* **(+) more control**: feeling more confident in ability to reduce time, checking, or interactions **(-) less control**: feeling less confident in ability to reduce time, checking, or interactions	
	Teens	Adults	Teens	Adults
Reels	(-) time spent; share of time spent	(-) time spent, share of time spent, engagements, and passive time		
Stories	(-) engagement	(-) time	(-) engagement + time	(-) engagement + time
News Feed		(-) passive + time		
Feedback (reactions, comments)	(-) engagement + time	(-) engagement + time	(-) engagement + time	
Notifications	(-) sessions from push; notification checking from in app	(-) sessions from push; notification checking from in app	(-) sessions from push; shorter time to star session following push	(-) sessions from push
News	(+) engagement + time	(+) engagement + time	(+)	(+)
Photos	(+) time + share of time spent	(+) time + share of time spent	(+) time + share of time spent	(+) time + share of time spent
Groups				(+)

*Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation

Facebook data dictionary and more details

29 duct details	Predictor details	Attention Strength of associations		Capability Strength of associations	
		Teens	Adults	Teens	Adults
Reels	~50% of users watch reels (teens 49% and adults 55%) with ~20% of users surpassing 1-hour weekly	(-) Total time and share of ts on reels. Total time on reels is ~10th most important feature, and it has ~10th highest correlation (2nd out of surfaces), while it has the strongest marginal effect from linear regression	(-) Total time spent, share of time spent, engagements (reactions, comments, and reshares) and passive time all surface. Reels fall below stories and passive news feed use in importance. Correlations and feature importance is anywhere from ¼ - ½ the most important features		
Stories	Stories engagement includes reactions, replies, reshares, and creation. Engagements could be on others stories or in response to a reply/reaction to your own. ~45% of users are spending time viewing stories, with 30% engaging with stories. While similar share of teen and adult view/engage with stories, teens who engage do so at well over double the volume averaging ~30 stories engagements per week.	(-) Stories engagement and engagement rate per time spent. In a top tier of important features with feed, notifs, and photos. Additionally having one of the strongest correlations	(-) Time spent on stories has the largest marginal effect from regression estimates (30 minutes L7 on stories is associated with 0.4 lower response on 1-5 attention scale). As well as being one of the strongest correlates out of surfaces and showing up in terms of both total time and share of time in the top 10 for feature importance	(-) engagement + time 2nd most important feature and second largest marginal effect from regression (30 min on stories L7 associated with 0.2 lower capability composite)	(-) engagement + time Strongest negative correlate and the 2nd & 3rd most important features. Also the second largest marginal effect (30 min on stories L7 associated with 0.6 lower capability composite)
Feedback (reactions, comments)	In survey 60% of respondents spend time looking at comments and 50% engage with comments. Adults have 5%-7% higher rates of both than teens. Adults are spending on average 15 min consuming or engaging with comments while teens are at 7 min. 13% of adults and 6% of teens are spending >30 minutes in comments per week. 11% of adults and 9% of youth have 14+ engagements in comment section weekly.	(-) engagements with (likes, replies, etc..) and time spent in comments is related to low attention use in both teens and adults. These are in the 10-20 most important features with ~1/3rd the correlation and 1/3rd the marginal effect on low attention response as our most important features like notifications or stories.		(-) engagement with comments and time spent in comments show up at ~⅓ - ⅕ of the marginal effect and variable importance as the most important features.	

Slide 88

28 "Feature importance" -- relative strength of contribution to predicting the likelihood of low control
Elena Davis, 12/22/2022

27 some operationalization here
Elena Davis, 12/22/2022

29 (aka 'gain in predictive accuracy')
--while controlling for everything else
Elena Davis, 12/22/2022

Continued: Facebook data dictionary and more details

32 duct details 33	Predictor details	Attention Strength of associations		Capability Strength of associations	
		Teens	Adults	Teens	Adults
News Feed			(-) Feed time is 2nd most important feature, with share of time also in top 10. Passive time is the main factor.		
Notifications	While low attention is related to both push notification initiated sessions and checking from within the app, we see capability being largely related to just push notification initiated sessions.	(-) Push notif sessions have the strongest marginal effect with a 1-standard deviation increase in push notif sessions being associated with 5.5%-5.75% higher rate of low attention use for youth and adults respectively. We also see notification checking from within app associated with low attention being the 5th-7th most important feature in predicting low attention and a correlation of -0.10 (for both youth and adults).		(-) Push notif sessions has the strongest marginal effect from regression analysis, with 1.0%-1.3% increase in low capability in association with 1 standard deviation increase in push initiated sessions. We also see the number of push notif sessions as the most correlated with the capability composite (a negative correlation). This is consistent for youth and adults	
News	In the survey sample news tab usage is pretty low, less than 2% of respondents were WAU of news tab and ~1% had 10 minutes average time spent on News.	(+) across both teens and adults and attention and capabilities composites for control that time spent on news tab and engagement with content in news tab has the largest marginal effect with respect to more positive responses. We see ten minutes L7 on news tab is associated with a full point higher response on the 5-point attention score. On the capability 5-point composite score we see 10 minutes L7 associated with 0.5-0.7 higher feeling of capability for adults and teens respectively. These news users are spending slightly more time than the non-news users (if we look at respondent with >30 min total time spent L7) the news users average ~15 more minutes on FB per day (10 of this is news). They are spending 15 more minutes combined on News and Groups, plus 20 more minutes on Feed, while spending ~15 minutes less on Reels + stories (reels is 2/3rds of that).			
Photos	2/3rds of users surveyed are spending time on photos consistent across age. With 30% spending greater than 1 hour weekly. We see engagement rates pretty similar at ~1/3rd for youth and adults and ~3 photo engagements per user.	(+) Across both teens and adults we see spending more time on photos and a higher share of time spent on photos is related to better feelings about low attention use and low capabilities. They come out as the top feature in terms of variable importance for youth and adult attention as well as youth capability, and the 5th most important feature for adult capability. When looking at the marginal effects they are close to as large on the positive side as stories are on the negative side, for attention and youth capability. When we look at users within overall time use buckets (10 min per day buckets) we see that users with >10 percent of time spent in photos in those groups have 10% lower rates of low attention and five percent lower rates of low capability than their counterparts (looking at users between 30 min - 2 hour daily time spent we see the 10%+ photo users spending 15 more minutes daily on photos). Above a certain amount of time spent (2 hours daily) these higher control results go away.			
Groups	9% of adults and 4.5% of teens in the survey spent more than 10 minutes in groups for the week. The engagement rates come in at 12% of adults and 8% of teens having engaged with groups content (reacted, comment, or posted in group) in the last week.				(+) time spent in groups is the 7th most importance feature and has a feature importance between ½ - ⅔ the most important features like Notifications, News and Stories. In regressions we see ¼ -⅓ the marginal effect as News and stories.

Slide 89

31 "Feature importance" -- relative strength of contribution to predicting the likelihood of low control
Elena Davis, 12/22/2022

30 some operationalization here
Elena Davis, 12/22/2022

32 (aka 'gain in predictive accuracy')
--while controlling for everything else
Elena Davis, 12/22/2022

33 wanting to note here to capture from above --your comment 'The stronger piece is push notification initiated sessions, that's tied with time spent as the strongest correlate. From the regression analysis we see that the most significant coefficient is the one on the count of push notif sessions."
Elena Davis, 12/23/2022

13 oduct	Additional findings/interpretation
Reels (-) attention	Experts have pointed to concepts like mindless scrolling as well as platform affordances such as recommendations, ephemeral content, and autoplay associated with low control (report). Activity on ephemeral surfaces like Stories might be related to not wanting to miss something, and wanting to respond in a timely way to others (expert feedback). Common critiques of excessive time spent, or passive time spent that might displace other meaningful activities, implicate endless feed or long periods of content consumption (Liu et al. 2019 meta-analysis). Browsing limitless content, along with intermittent reinforcement of seeing high-value content, has been proposed as al mechanism of continuing to spend time on social media despite intentions (example cites 1, 2, 3). Past PU research (notably that did not assess low attention specifically) did not find proportion of time spent on News Feed as associated with PU (Cheng et al. 2019): *"Overall, our findings on messaging activity and time spent contradict an image of people experiencing problematic use because of hours of unintentional scrolling or serially watching videos. Instead, they paint a picture of people spending more time browsing profiles and messaging others."* With a broader conceptualization of low control, these present findings show Reels + NF as related to low attention cluster specifically. Though nuanced, passive consumption experiences overall may not contribute as positively to well-being as more active or interactive experiences that support connecting with others (see connections literature review, emotional impact literature review).
Stories (-) attention & capability	
News Feed (-) attention [adults]	
Feedback (reactions, comments) (-) attention & capability [teens]	Lightweight actions were more bundled in with passive use in terms of association w well-being (?) (Burke et al. 2016)
Notifications (-) capability [adults]	Notifications – (experts) (Cheng et al. 2019).
News (+) attention & capability	Potentially related to more intentional consumption because switching tabs and/or deliberately navigating to another surface. Different subsamples of users who commonly engage in these surfaces 1. News users tend to spend more time on the app (~15 more min daily), however that is driven by news tab, groups and news feed time spent which they have a combined 35 mins more of daily and less time on reels and stories (~15 mins less). 2. Photos positive associations are related to users that spend 10% or more of their time viewing photos (~7% of survey respondents were above 10% and 4% were >10% and >10 min on photos). We see for these photos users that they spend about 30 min less per week with the gap driven by 1-hour less time on video and 30 mins more on photos 3. Groups users behave a bit differently than the non-groups users in our survey sample. They were far more likely to engage with news, 3.5x the engagement rate, and spent less time consuming video (~15 min less a week). Adults in our sample have 2x the
Photos (+) attention & capability [teens]	
Groups (+) capability [adults]	

Slide 90

12 I think this could also be related to photos not having chaining, so it's harder to get into repetitive/automatic consumption.

Photos in stories or photos that have been pushed together into a reel that is multiple photos wouldn't be labelled as photo surface (instead reels or stories surface) so that edge case where photos can be chained wouldn't be included in our photo surface.
Josh [redacted] 12/28/2022

13 I added comments in prior slide but we see within overall time spent buckets users with more photos time spent (not just a higher share for lower ts people) have markedly lower low attention and low capability scores (10% better on attention and 5% better on capability).

for example in the big >30 min daily time spent cohort (results are similar if we bracket more granularly) users with >10 min daily on photos or >10% ts on photos have 40% bottom 2 box attention compared with 53% bottom 2 box for the rest. Capability is around 15% and 20% for those two groups respectively.
Josh [redacted] 12/28/2022

Instagram usage (non-causal) associations with perceived control

Product	Attention *More time spent within these products associated* with:* **(+) more control**: lower frequency of losing track of time and/or habitual use **(-) less control**: higher frequency of losing track of time and/or habitual use		Capability *More time spent within these products associated* with:* **(+) more control**: feeling more confident in ability to reduce time, checking, or interactions **(-) less control**: feeling less confident in ability to reduce time, checking, or interactions	
	Teens	Adults	Teens	Adults
Explore tab	(-) passive + time	(-) passive + time	(-) passive + time	(-) passive + time
Clips (Reels tab)	(-) passive + time	(-) passive + time		
Video consumption (any surface)	(-) frequency of chaining + time	(-) frequency of chaining + time		(-) video chaining
Home	(+) share of time spent (-) total time and passive time spent	(-) passive + time	(-) passive + time	
Profile		(-) passive	(-) engagement + time	
Stories	(-) creation		(-) creation	(-) creation
Direct	(-) long dm sessions	(-) long dm sessions	(-) engagement + time	(-) engagement, passive, + time
Notifications				

91

*Associations were evaluated 3 ways to highlight converging signals: unique predictors in a multivariate regression, strongest contributors in a predictive decision trees (xgboost), and correlation

Slide 91

16 Clips is the clips/reels tab, explore is explore tab, and video is any video consumption across any surface.

For all three we see chaining depth as an important negative factor.

For chaining it appears total volume of chaining impressions is more important than how deep into chaining you get. So frequently getting a few videos deep into chaining is the more important factor than the occasional deep chaining (e.g. depth of 10-20+).

Daily max depth of chaining is less than half as important of a feature as total impressions in chaining. We see a 1-stddev higher number of impressions in chaining associated with 4% higher probability of being bottom 2 box, while a 1-stddev higher depth of chaining is under 2%.
Josh [redacted] 12/22/2022

18 20% of overall DAU producer stories and almost 30% of teen DAU produce stories.
Josh [redacted] 12/22/2022

19 I think this is capturing users that just come, check in on whats new on their feed and then leave without using DM's, explore, clips, etc.. They likely spend less time. I can check this quick.
Josh [redacted] 12/22/2022

20 So home overall ts and home share of ts are less correlated than share and total ts on other surfaces. For example clips, dm and explore TS and Share of TS have 0.6-0.7 correlation while on home it's < 0.3.

We also see share of TS with negative 0.23 - 0.3 correlation with time spent on explore, clips, and dm.
Josh [redacted] 12/22/2022

36 yup so important. so it's chaining AT ALL -- which actually feeds into the autoplay issue
Elena Davis, 12/23/2022

37 whereas watching a video and having a pause before starting the next, or opting in to an autoplay experience within the context of that specific viewing instance, makes sense as a solution
Elena Davis, 12/23/2022

35 Jess question - how clips and video specified and differentiated?
Elena Davis, 12/23/2022

38 love this actually for deep immersive consumption possibly being intentional and something we can support (IG reels interested here)
Elena Davis, 12/23/2022

17 Also Feed creation so like creating posts. The stories effects are bigger and more significant as well as stories creation being a far more important feature.

I believe this is do to frequency of creating stories is higher than feed. For reference stories producer dau is >200 million compared to 30 million for

feed producer dau.
Josh [redacted] 12/23/2022

39 yup so what I'm hearing is we can understand how posting content contributes to downstream difficulties with control, across formats. but stories would make more sense to start bc of prevalence of the behavior. also - I still wonder if there's something here re: ephemerality possibly creating more time-bound pressure
Elena Davis, 12/23/2022

34 is it right to indicate 'frequency of chaining' here and possibly for adults too?
Elena Davis, 12/23/2022

15 Yes for both adults and youth on attention you could cite frequency of video chaining.

Video TS overall and share of time on video are important, a lot of it is related to chaining, explore and clips tab though. For example the correlation with all those and video ts is 0.3-0.6
Josh [redacted] 12/23/2022

14 Need to investigate more, this result is weird when diving in
Josh [redacted] 1/12/2023

Instagram data dictionary and more details

43 duct details	Predictor details	Attention Strength of associations		Capability Strength of associations	
22					
23		Teens	Adults	Teens	Adults
Explore	For Teens and Adults ~⅓ had >10 mins L7 on explore, ⅙ had >60 mins L7 and 1/12 had > 2 hrs on explore L7. We also see passive behaviors (scrolling, swiping, etc..) related to low attention use.	(-) passive consumption behaviors like scrolling, swiping, and tapping into fullscreen/chaining and spending more time in explore are one of the strongest marginal effects from regression. We see 1-std-dev increase on both passive and time spent associated with 8%-10% higher low attention usage. The The results from gradient boosting feature importance and correlation analysis are more modest, in the 10-20th strongest associations.		(-) share of IG time spent on explore, passive behaviors on explore and time spent on explore are the predictors related to low capability. We see the associations are weaker with respect to capability than attention (marginal effects are ⅓ - ½ of he size of those on attention as well as being smaller than story/feed creation, notifications, dm's, passive home feed, and profile. Feature importance and correlations are between ¼ - ⅛ of the importance of the surfaces mentioned immediately above. Unlike the attention piece where we see some interactive effects of both above median engagements and passive behaviors being particularly associated with low attention, we don't see any such association with engagements on capability, it's just a passive story on capability.	
Clips (Reels tab)	11% of teens and 6% of adults in the survey sample spent >1 hour on clips tab weekly. On average we see the teen sample wit 32 minutes on clips tab weekly compared with 14 min for adults. Among the ¼ of teens and 1/7 of adults that have >10 min of clips TS weekly, the average clips tab ts is 2:20 for teens and 1:35 for adults.	(-) For both adults and teens we see the correlations and predictor feature importance associated with clips tab time spent and passive clips tab usage rate as only modestly important, coming n at less than half the importance and correlation as explore, profile, direct, and passive home usage. We see stronger correlations with clips tab for teens than adults but still ~50% of the importance and correlations of the most important features. However, clips tab has strong marginal effects on rates of low attention usage, coming in 1 percentage point lower than the explore tab results (which are the largest for a surface). We see for passive usage that once we restrict to users above 60 minutes of clips tab weekly being above median on passive usage rate (scrolling, swiping, etc..) is associated with a 13 percentage point higher rate of perceived low attention usage.			
Video	90% of the users in our survey spend >10% of their time on IG on video, while 67% spend 30% or more of their time on video and 40% spend more than half of time on video. Half of the survey spent >1 hour on video content a week with the average time on video nearly 2.5 hours weekly. The users averaged getting into video chaining 65 times in the last week with an average depth of ~3.5 videos viewed per chaining session. With 13% of users getting into chaining 175+ in a week (25x daily)	(-) For both adults and teens we see share of time spent watching video and overall time spent on video in the 10 most important features. Video ts is also the second most correlated surface with low attention usage. Video chaining especially the number of times getting into video chaining was associated with low attention usage. The number of chaining vpv's has 1.6x higher marginal effect and higher feature importance as well as 1.7x stronger correlation with low attention usage when compared to average chaining depth. We see the number of times entering video chaining with a correlation of -0.11 to low attention and -0.06 to low capability which is 2x-3x stronger than average max depth of chaining session (how many videos deep a user gets on average).			(-) frequency of getting into video chaining and timespent in video chaining are the second strongest correlates with low capability for adults. While we see frequency of chaining and total number of chaining impressions related to low capability the average max depth gone into chaining isn't related to low capability. It seems the primary relation to low capability is frequently getting into chaining perhaps unintentionally being driven into chaining when viewing a video.

Slide 92

21 Clips is the clips/reels tab, explore is explore tab, and video is any video consumption across any surface.

For all three we see chaining depth as an important negative factor.

For chaining it appears total volume of chaining impressions is more important than how deep into chaining you get. So frequently getting a few videos deep into chaining is the more important factor than the occasional deep chaining (e.g. depth of 10-20+).

Daily max depth of chaining is less than half as important of a feature as total impressions in chaining. We see a 1-stddev higher number of impressions in chaining associated with 4% higher probability of being bottom 2 box, while a 1-stddev higher depth of chaining is under 2%.
Josh [redacted], 12/22/2022

41 yup so important. so it's chaining AT ALL -- which actually feeds into the autoplay issue
Elena Davis, 12/23/2022

42 whereas watching a video and having a pause before starting the next, or opting in to an autoplay experience within the context of that specific viewing instance, makes sense as a solution
Elena Davis, 12/23/2022

40 Jess question - how clips and video specified and differentiated?
Elena Davis, 12/23/2022

43 love this actually for deep immersive consumption possibly being intentional and something we can support (IG reels interested here)
Elena Davis, 12/23/2022

22 When drilling into engagements (reactions, comments, reshares, etc..) and passive behaviors (scroll, etc..) we see that passive behaviors further strengthen the relationship.

For example if we split users at ~median passive behaviors per minute on explore tab, among users with >60 min of explore TS L7 the above median passive actions have nearly 9 percentage point higher low attention, for users with 10-60 min L7 on explore that same gap is ~8 percentage points with more passive group having again the higher rate.

When we do this same thing for active engagements that delta is less than half (1.5% - 3%). There does seem to be some additive effect though whei the users above the median on both passive behaviors and engagements have 12-15 percentage point higher rates.
Josh [redacted] 12/29/2022

23 unintentionally getting into chaining
Josh [redacted] 1/5/2023

Instagram data dictionary and more details… Continued

24 duct 25 ails	Predictor details	Attention Strength of associations		Capability Strength of associations	
		Teens	Adults	Teens	Adults
Home	Average time spent on feed is ~2 hours weekly among the sample, 20% of users are surpassing 3.5 hours weekly (or 30 min daily on feed)	(+/-) Higher share of time spent on feed is correlated with higher attention usage, We see that users that have 50% or more of time spent on home feed (~1/3rd of teens) have 5% lower rates of low attention usage, they have 90% lower time spent on explore and clips (10 minutes weekly compared to > 2 hours on average weekly). They also have less usage of direct messaging and lower stories creation (50%-75% lower). Share of time on home feed is the 6th most important predictor and has the largest positive correlation with high attention usage. While spending more time on feed overall, especially with lower engagement rate and high passive behavior rate (swiping and scrolling) have strong negative relationships with attention with both in the top 5 most important predictors of low attention usage for teens. Passivity only becomes important once we get above an hour per week spent on home feed, below that we see no relationship, above 1 hour we see that 3 additional passive behaviors (scroll, swipe, etc..) is associated with 6% higher rates of reporting low attention usage.	(-) Spending more time on feed overall, especially with lower engagement rate and high passive behavior rate (swiping and scrolling) have strong negative relationships with attention with both in the top 5 most important predictors of low attention usage for teens. Passivity only becomes important once we get above an hour per week spent on home feed, below that we see no relationship, above 1 hour we see that 3 additional passive behaviors per min scroll, swipe, etc..) is associated with 4% higher rates of reporting low attention usage. (for users with over a minute of feed time spent Median passive behaviors per min is 3, mean is 6 and std. dev is 9)	(-) A higher amount of passive time spent on feed is We see passive action and the rate of passive actions per minute of feed time as the 2nd and 8th most important features.	
Profile	20% of teens and 24% of adults are spending an hour per week on profile.		(-) For adults we see time spent on profile as the most important feature for predicting low attention usage, while it's a behavior not relevant to a large share of users - it only factors into low attention usage prediction for 5% of users - it drives 21% of the predictive accuracy overall. Profile TS also has the most significant estimate from regression analysis where we see a 1-standard deviation shift (10 min daily on profile) associated with 7% more low attention usage construct. We see it related to passive behaviors on profile (consumption e.g. scrolling, swiping, etc..) and not making changes or interacting with things in profile. Passive behaviors is the 4th most important feature. Both passive behaviors in profile and profile ts are some of the strongest correlates (-0.15 and -0.18 respectively). When we look at the combination of both a 1-standard deviation shift in time spent on profile and 1-standard deviation shift in passive behaviors on profile we see it associated with 17% higher rates of low attention usage	(-) Profile engagements (like, comment, post, change, etc..) are the third most important predictor of low capability to control, with time spent on profile as the fourth most important feature. Combined these two behaviors drive 8% of the predictive accuracy. A 1 standard deviation increase in time spent on profile (60 minutes per week) is associated with an almost 5% increase in reporting low control capability. It has the second largest marginal effect after storied creation.	

Slide 93

24 Need to double check, I'm almost certain this is per min, since three passive actions in aggregate is nothing
Josh [redacted] 1/11/2023

25 Need to dig into this more. It's not making a lot of sense when I get into some of the cuts
Josh [redacted] 1/12/2023

Instagram data dictionary and more details… Continued

Product details	Predictor details	**Attention** Strength of associations		**Capability** Strength of associations	
		Teens	Adults	Teens	Adults
Stories	Creation	(-) Stories creation is the third most important feature predicting low attention usage. It is responsible for 6% of the accuracy gain in predicting low attention usage. We also see it as one of the strongest correlates having a -0.14 correlation with attention usage construct. We see an additional 10 stories posted associated with 3% higher rates of low attention usage among teens. Overall the 20% of survey respondents that had 10 or more stories created weekly had 8% higher rates of low attention usage.		(-) Stories creation is the strongest correlate with capability to control with -0.14, stronger than even the global behaviors like time spent or sessions count. Stories creation is the most important feature in predicting low capability to control contributing 18% of the predictive accuracy for youth and 8% for adults. Stories creation has a consistent relationship with low capability to control, with us seeing ~1.7% higher rates of reporting low capability to control for each additional 10 stories created, with the 20% of survey respondents that posted 10+ stories in the week having ~5% higher rates. We similarly see feed creation related to low capability to control, with survey respondents that created 4 or more feed posts having 6% higher rates of low capability to control. Each additional feed creation is associated with 0.3% higher rates of low control (the 2% of users posting 10 or more times on feed weekly have 11 percentage point higher low capability 28% vs. 17%)	
Direct	Engagement, passive,	(-) Long dm sessions - sessions with >10 minutes of time spent in DMs - are associated with low attention usage for both teens and adults. 11% of adults and 14% of teens have >=1 long DM session a day - 4% of adults sessions and 6% of teen sessions for the survey respondents are long dm sessions. We see that having 7 additional long dm sessions is associated with 3.5% higher rates of low attention usage for both adults and teens. A 1-standard deviation increase is coming in at ~½ the marginal effect of the more prominently associated surfaces (explore, clips, video, home, and stories creation). We don't see DM's surface in other ways in a strong way in association with low attention usage being outside the top 15 predictors for both age buckets, and having more modest correlations between -0.08-0.10 which is outside the top 5-10 highest correlates.		(-) For both adults and teens we see time spent on DMs as the 7th strongest predictor of low capability to control and 4th strongest unique surface predictor (behind only profile, stories creation and passive home). We see DM's contributing 7% of our predictive accuracy for both age buckets. For both teens and adults direct TS spent with correlation of -0.08 are between the 5th and 10th most strongly related to control capability. A 1 standard deviation increase in direct usage whether time spent (20 min daily), engagements (3 engagements daily), sessions (10 daily) is associated with a 2%-3% higher rate of low control capability. After control for notifications which are increased as a result of DM usage we still see half of the marginal effect remain with 1%-1.5% higher low control capability exclude the impact of notifications.	
Notifications		(-) For youth notifications are the 20th most important feature while for adults it's 15th most important. For both age cohorts we see similar negative 0.1 - 0.11 correlation with attention, this is between 5th-10th strongest correlations for surfaces and 10th-15th when including global behaviors (like time spent, sessions, and time of day). Notifications (time spent on activity, notif impressions, notif clicks, push initiated sessions) all have between 3%-3.5% higher rates of low attention usage associated with 1-standard deviation increase (~60 imps and 20 clicks) effect sizes are closer to 3% for youth and for adults closer to 3.5%. This makes notification the 7th largest marginal effects for surfaces. The relationship between push notifications and Low attention usage is almost entirely mediated by stories/feed creation and DMs, with 75%-95% of the association between notifs and low attention usage being explained by the users DM/Creation behaviors (75% for adults and 95% for youth). For Adults we see notification impressions being far more important than clicking through or initiating a session from a push notification (2/3rd of the relationship is explained by impressions)		(-) For teens notifications are the 11th most important feature and for adults notifications are the 9th most important feature. For both age buckets we see notifications as the 3rd-4th most correlated surface with capability to control and between 5th-10th when including global behaviors. We see the marginal effect of notifications on capability to control coming in a little below DM's with a 1-standard deviation increase in notification imps (60 weekly) or clicks (20 weekly) associated with 2% increase in low control capability, unlike low attention we see the notifications having an impact on low capability to control even after controlling for the effects of stories/feed creation and DM. One half of the relationship is explained by notifications alone (1%). I think notifications also deserve a callout in the low capability 3rd tier. We see time spent on activity in 9th (teens) and 11th (adults) most important feature. We also see Sessions originating from push notifications (aka clicks) coming in just a little below the marginal effect of direct time spent (2% vs. 2.5% relative to 1 std. dev. increase).	

Instagram usage (non-causal) associations with perceived control

26oduct	Attention (+) more control (-) less control	Capability (+) more control (-) less control	
Explore	(-) passive + time	(-) passive + time	Similar as Facebook slide: • Browsing limitless content • Autoplay • Recommendations Home also parallels with passive browsing experiences; Positive association for teens is related to having a higher share of TS on home. We see that users that have 50% or more of time spent on home feed (~1/3rd of teens) have 5% lower rates of low attention usage, they have 90% lower time spent on explore and clips (10 minutes weekly compared to > 2 hours on average weekly). They also have less usage of direct messaging and lower stories creation (50%-75% lower). They spend only slightly less time on video (within 20% of video ts) however the teens that spend more time on home get into chaining less than half as many times and spend only 1/3rd the amount of time in chaining. • Friend content is larger share of teen consumption (27% vs. 17% for genpop) and home is more concentrated with friend content than other major surfaces • Friend content is consumed earlier in the home feed / in sessions, so longer passive browsing will get further away from the initial friend content time spent on home (first third of sessions have 25% higher friend content) [link]
Clips	(-) passive + time		
Video	(-) Frequency and volume of chaining + time	(-) Frequency of chaining [adults]	
Home	(+) Share of time [teens] (-) passive + time [teens & adults]	(-) passive + time [teens]	
Profile	(-) passive [adults]	(-) engagement + time [teens]	Specifically found in (Cheng et al. 2019) - unclear the function of browsing profile; past work found no differences in browsing own vs others' profiles, or activity on profiles. Profile may be related to revisiting own past content; social curation, editing/updating (cite; cite; cite)
Stories	(-) creation [teens]	(-) creation	Creation may beget more responses and thus need to continue to use, all in a time-bound ephemeral way Also possible to lose track of time during creation, such as seen in certain cohorts using filters/perfecting their posts/curating (>? Need to find research) Stories creation is high intent (cite)
Direct		(-) engagement + time (-) engagement, passive, + time (+ notifs?) [adults]	Parallels with notification findings from Facebook

Slide 95

44 This is Reels
Elena Davis, 12/22/2022

26 Clips tab / the reels tab on IG. A reel could also be watched on explore so clips isn't all reels watching but watching reels from clips tab or chaining of of clips tab.

Video would encompass all reels - and other video - on any surface.
Josh [redacted] 12/22/2022

Appendix: FAQs

FAQ: How were the survey items selected?

Details of Approach

Focus is on measuring users' perception of control over their time and experiences on our platforms (see note)

Measurement review of PU constructs & existing survey items

Pre-existing measures of PU had few items on control specifically.
A broader set of control items were developed and variants were written to cover control over different aspects of platform use

August

Lit review
Past surveys 1, 2
Measurement audit + takeaways
List of items

Pilot test potential survey items covering key control constructs

Off-platform quantitative pilot test in 6 markets & cognitive testing in 2 markets, with both teens and adult FB & IG users

September - October

Research plan
Cog testing results
Pilot results + notebook
All items tested

Refine survey items and field on-platform measurement survey

On-platform fielding in 12 markets, to ~120K teens and adult users across FB & IG.

November - December

Research plan
Simon FB
Simon IG

Specify constructs for 'sense of control'

Factor analysis and invariance + reliability testing to define composite variables:
low attentiveness + low capability

December

Results
Analysis notebook

Identify on-platform correlates of low control

Identify product surfaces to prioritize where to focus PU solutions, and behavioral patterns to develop next phase of PU heuristics

December - January

Survey item development (snapshot)

26 survey items were developed to cover a range of aspects of control, and control applied to several areas of behavior. (Sources: PU measurement audits and review 1,2,3). These were all included in an off-platform pilot survey (research plan)

Selection of survey items to illustrate coverage	Time spent	Checking
Awareness	How aware are you of the amount of time you spend using [Instagram]/[Facebook]?	How aware are you of how often you check [Instagram]/[Facebook]? *Checking includes opening the app, checking a notification, and refreshing your feed for updates.*
Attentiveness	How often do you find yourself spending more time than you wanted to without realizing it, when using [Instagram]/ [Facebook]?	How often do you find yourself checking [Instagram]/ [Facebook] automatically without realizing it?
Capability (self-efficacy)	How confident are you that you would be able to reduce the amount of time you spend using [Instagram] /[Facebook], if you wanted to?	How confident are you that you would be able to change how often you check [Instagram]/ [Facebook], if you wanted to?
Locus of control	How much is the decision about how much time you spend on [Instagram]/[Facebook] your own decision?	Are there external reasons, apart from your own decisions, that impact how often you go on [Instagram]/[Facebook]? *This could include other people determining how much you can access the app, or needing to use the app for school or work purposes.*

Cognitive testing (snapshot)

A subset of items were cognitively tested among teens and adults in the US and India; see full report and plan

Summary of Item Performance

No changes recommended
Minor changes recommended
Substantial changes recommended

Item	Summary of Performance	Rating
Q4. How often do you find yourself spending more time than you wanted to without realizing it, when using [IG/FB]? Q4-ALT. How often do you find yourself losing track of how much time you spend on [IG/FB]? *a. Very often / b. Often / c. Sometimes /d. Rarely / e. Never*	• All Ps understood that this question was asking about the frequency at which they "lost track of time" while using Facebook or Instagram. Other terms that participants used to discuss this question were, "getting carried away," "getting sucked in," or "spending more time than intended." • Most Ps thought Q4 and Q4-ALT to be asking about similar concepts. • We recommend Q4 because of challenges in the Hindi wording in Q4-ALT.	
Q5. How often do you use [IG/FB] to do something specific, like opening the app for a particular reason or wanting to accomplish something? *a. Very often / b. Often / c. Sometimes d. Rarely / e. Never*	• All Ps understood that this question was asking if they go onto IG or FB with a particular purpose or goal in mind. • Ps were split about whether "browsing" or "passing the time" counts as something specific. Conceptually, it's hard to distinguish these interpretations.	
Q6. When using [IG/FB] to so something specific, how often do you get distracted so you don't do what you originally wanted to? *a. Very often/ b. Often/ c. Sometimes/ d. Rarely / e. Never / f. I don't use [IG/FB] to do something specific*	• Ps mostly interpreted the question as intended, assessing whether they get diverted from their original purpose of using the app. • Two adult Ps considered specific tasks outside their use of IG/FB, such as not being able to complete work-related activities because they are distracted with the app.	
Q7. How much is the decision about how much time you spend on [IG/FB] your own decision? *a. Entirely my own decision/ b. Mostly my own decision/ c. Somewhat my own decision/ d. A little my own decision/ e. Not at all my own decision*	• Ps understood the question as intended, considering whether a person feels like it's their own individual choice how much they use the app or not. • There were differences on the factors that affected teens and adults' decision on the time spent using IG/FB. Teens mainly referenced parental control, while adults mentioned work related limitation. • External influences can increase or decrease time on the app.	

Pruning of survey items

Converging signals from cognitive testing and statistical performance in the quant testing led to final set of survey items, see process here.

Some specific findings:

- Questions asking about **the extent of control** (e.g., "How much control do you feel you have…?") had notable validity issues and previous localization concerns.
    - It is thus recommended to ask about control in different ways, with more of a focus on behavior and avoiding the use of the term 'control'
- Questions about **awareness of behavior** (e.g., "How aware are you…?) also had notable validity issues, and previous research has shown that people generally have inaccurate awareness of their app time spent (note) and these can have sources of bias
    - If a product aims to increase user awareness of specific behavior, consider evaluating awareness as self-evaluation of something specific (e.g.,'how often do you do xyz?' plus 'how confident are you in your answer?'/'how easy or difficult was it to answer that question?')

Final set of survey items and coverage of key aspects of control

	Control over time spent	Control over checking	Control over content consumption	Control over your interactions with others	Control over meeting goals *Items did not clearly assess accomplishing goals vs. getting distracted from them; lack of clarity around defining what a 'goal' or intention is for platform use*
Awareness *Not assessing awareness at a high level - validity concerns and known user inaccuracies*					
Attentiveness	1A	2A	3A	*Did not pilot a Q here - interactions more active so 'without realizing it' framing didn't work*	
Capability (self-efficacy)	1B	2B	*Exerting control over content becomes less about behavior and more about utilizing platform controls*	3B	
Locus of control	4B *Focusing on time spent to get at overall decision to use the platform*	*(Captured with 4B)*	*(Captured with 4B)*	*(Captured with 4B)*	

Pilot identified key control questions across two clusters

Pilot survey covered a range of aspects of control, applied to several areas of behavior. 7 key Qs emerged and were included in the on-platform survey along with an adapted question from the Main Tracking survey*

Low attentiveness around use (possible unintended, unaware, and/or habitual use)	**Perceived control capability** (perceived self-efficacy to exert control, if desired)
• How often do you find yourself spending more time than you wanted to without realizing it, when using Facebook/ Instagram? • How often do you find yourself checking Facebook/ Instagram out of habit, without realizing it? *Checking includes opening the app, checking a notification, or refreshing your feed for updates.* • How often do you find yourself getting caught up viewing a lot of content on Facebook/ Instagram without realizing it?	• How confident are you that you would be able to reduce the amount of time you spend using Facebook/ Instagram, if you wanted to? • How confident are you that you would be able to reduce how often you check Facebook/ Instagram, if you wanted to? • How confident are you that you would be able to stop yourself from interacting with people you don't want to on Facebook/ Instagram? *Interacting includes responding to other people through comments or direct messages.* • How much is the decision about how much time you spend on Facebook/ Instagram your own decision? ** ***this question ultimately excluded from final composite due to tracking a different construct (locus of control) and slightly poorer performance of CFA*

*"How often does your use of Facebook/ Instagram prevent you from fulfilling your responsibilities in other areas of your life (for example, work, school, or relationships)?"

FAQ: Do the survey items work for teens?

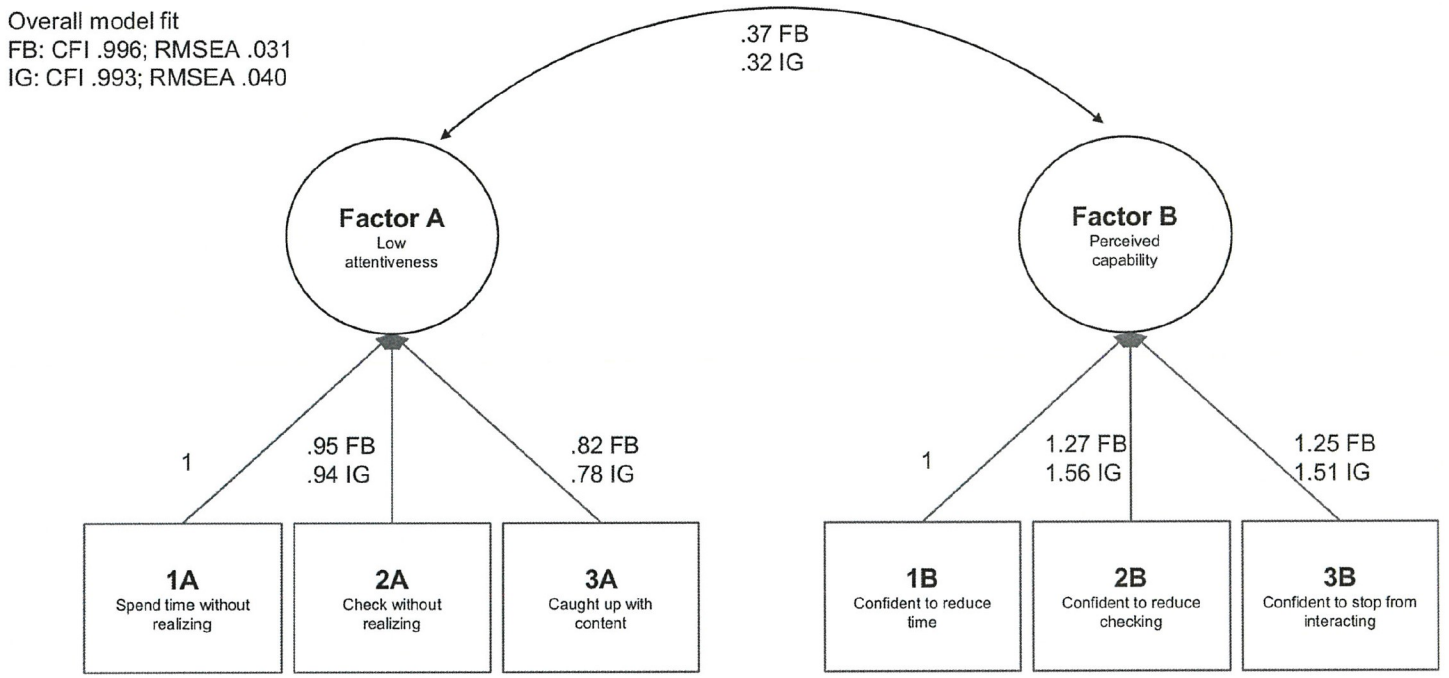
Cognitive testing validated that teens can answer these questions

Generally, teens and adults interpreted the questions similarly. While the responses and reasons for responses (e.g., teens referred to their parents) differed between teens and adults, their interpretation of the constructs being measured were similar.

- The items that were problematic and had validity issues for teens were removed from the final survey (e.g., 'how much control').
  - *Because I feel like a lot of people have different thoughts on how to define control because control can be a bunch of different things. (USA10 - teen IG)*
- Most Ps did not set a time that they wanted to be on the apps but they still felt as though they were spending more time than they should (based on other activities that needed to be done in the day).
  - *I try to manage according to my school work, home work, schedule of sleeping, food, so I try to manage Instagram time accordingly. (IND12 - teen IG)*
- Approximately one-third of Ps said their report on spending time would change if it they were asked if they did it "without realizing it"- there are times when they consciously decide to spend more time on FB or IG.
  - *If you are spending extra time and you are not realizing the time, then this thing happens often. However, if you are spending extra time but you are realizing it, then this thing happens sometimes. (IND09 - teen IG)*
- Most Ps referenced their self-control and their personal will of determining whether they use and spend time on the app or not.
  - *There's nothing really but myself that's controlling how much time I'm on it. No one's forcing me to go on it. I choose to go on it, and I choose to stay. I choose to stay on it. So I would say it's entirely my own decision. (USA13 - teen IG)*
  - *New trending reels or ... interesting content that attracts me most of the times, I don't feel like leaving it and I am always like 'let's see one more'. (IND12 - teen IG)*
  - *Mostly my own decision because my mom, she'll come in and be like, get off your phone when I'm doing something-- when I'm on my phone. (USA10 - teen IG)*

Survey items form 2 clusters, invariant for teens and adults

This means both teens and adults, across FB and IG, are answering the survey items in similar ways. We are consistently tapping into 2 clear clusters of sense of control, and can use these same items for everyone.



Consistent results from off-platform pilot survey (N=3750) across 6 countries (US, UK, IN, BR, KR, NG), youth and adults, FB and IG
CFA across entire sample strong fit (CFI = 0.986; RMSEA = 0.046); configural and metric invariance met across cohorts and platforms

FAQ: What are the ways to increase user control?

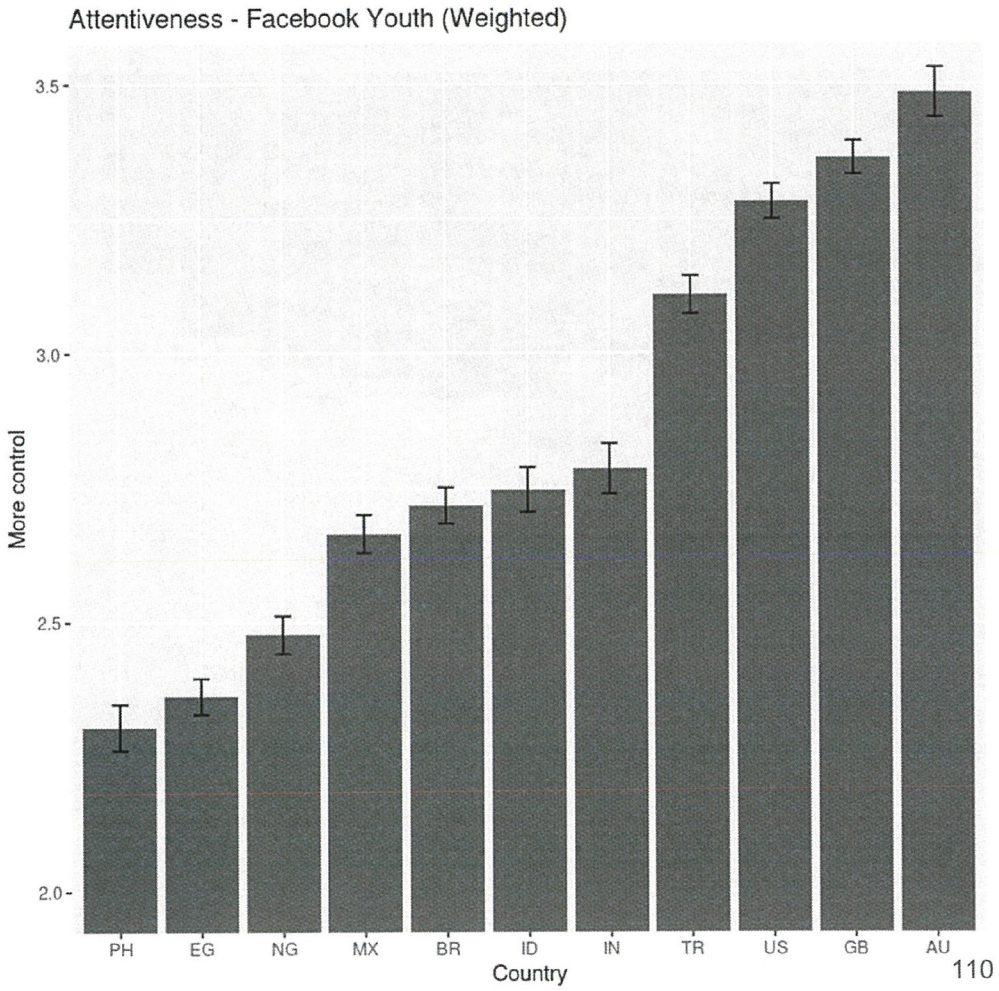
How can we increase user control, through products?

Control Levers[1]:	**Awareness** *Learn about and reflect on own behavior*	**Attentiveness** *Notice behavior in the moment*	**Intentionality & Choice** *Make explicit decisions about platform use*
What this might look like	Products that give people feedback about their own behavior and experiences on platform.	Products that give information in context - tied to specific behavior.	Products that enable behavior change or make change easier.
How it works	Information enables reflection about behavior and informed choices about desired changes.	Interrupts and draws attention to automatic behavior, slowing people down, and enabling more deliberate choices.	Helps people set goals, create limits, set explicit preferences.
Examples	Your Time on Facebook/Instagram	Feed Pause (FB) Take a Break (IG)	Personalized daily time limit

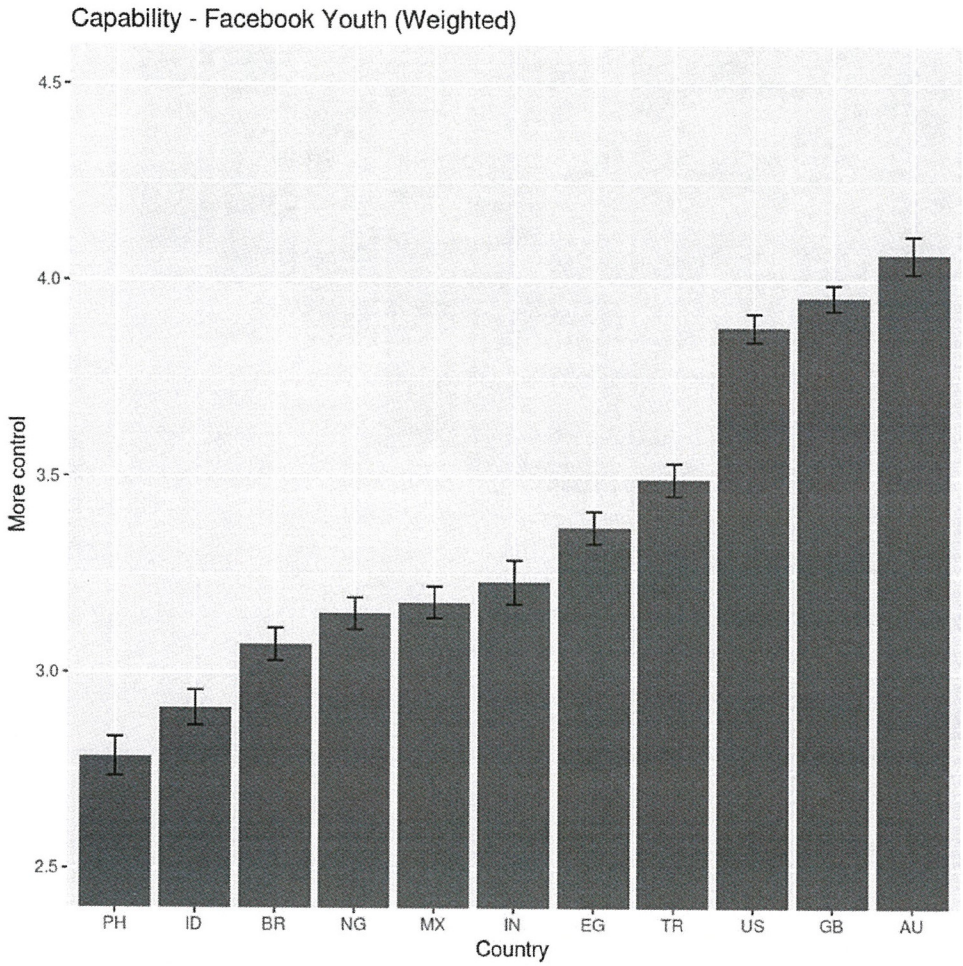
[1]*Levers are research-informed and expert-validated*

Additional country plots

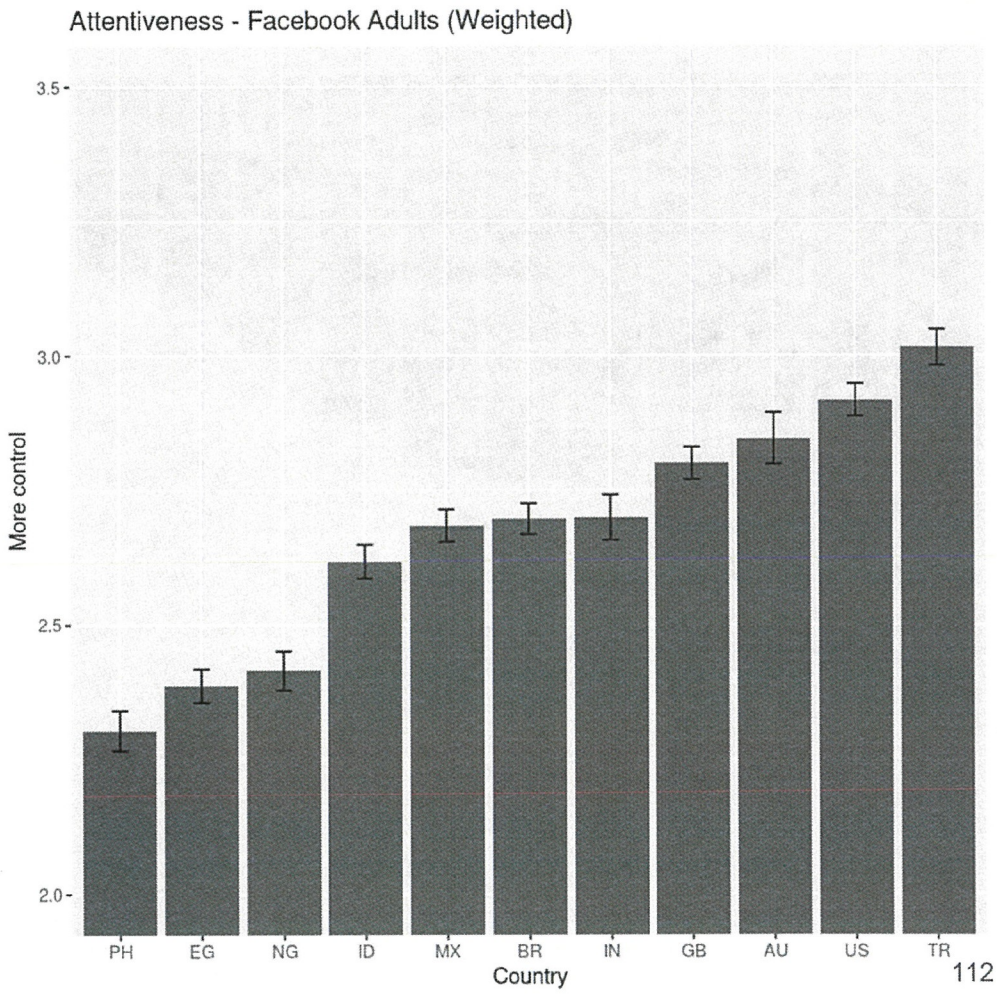
Facebook Youth: Mean perceived attentiveness by country



Facebook Youth: Mean perceived capability by country

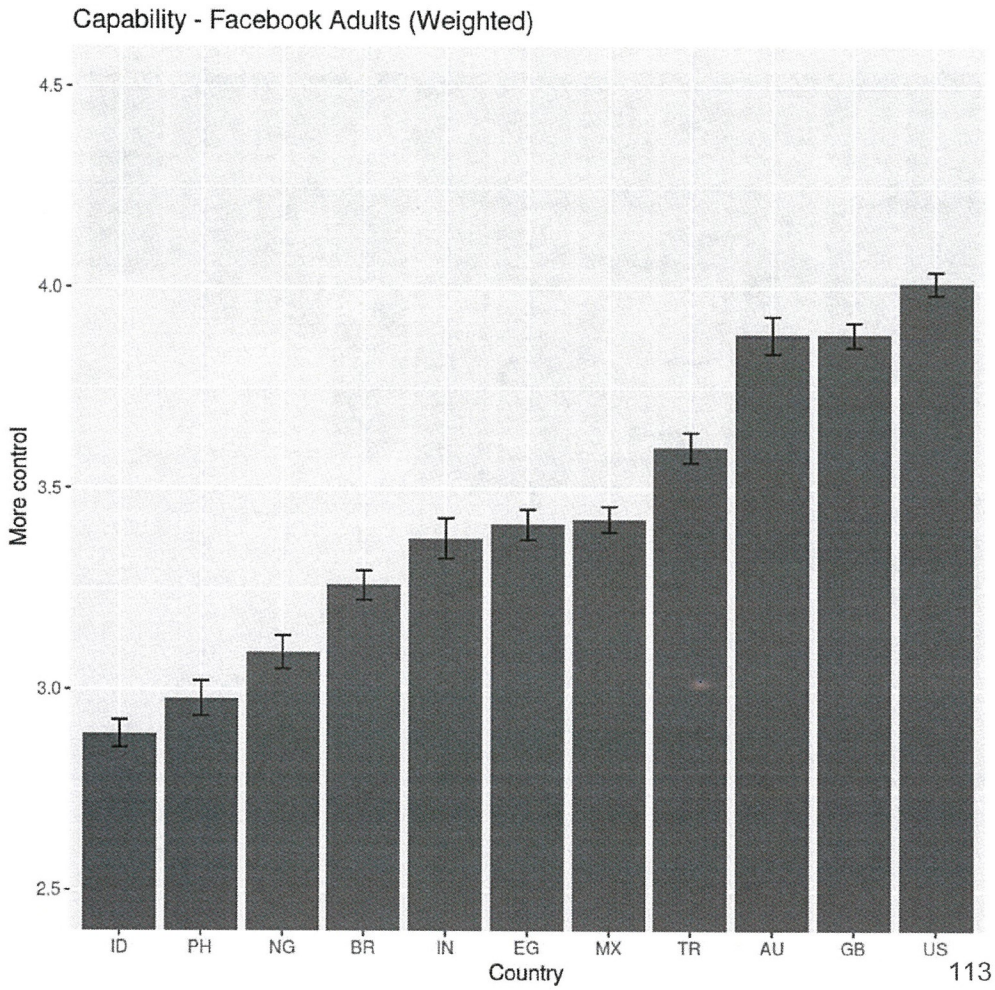


Facebook Adults: Mean perceived attentiveness by country

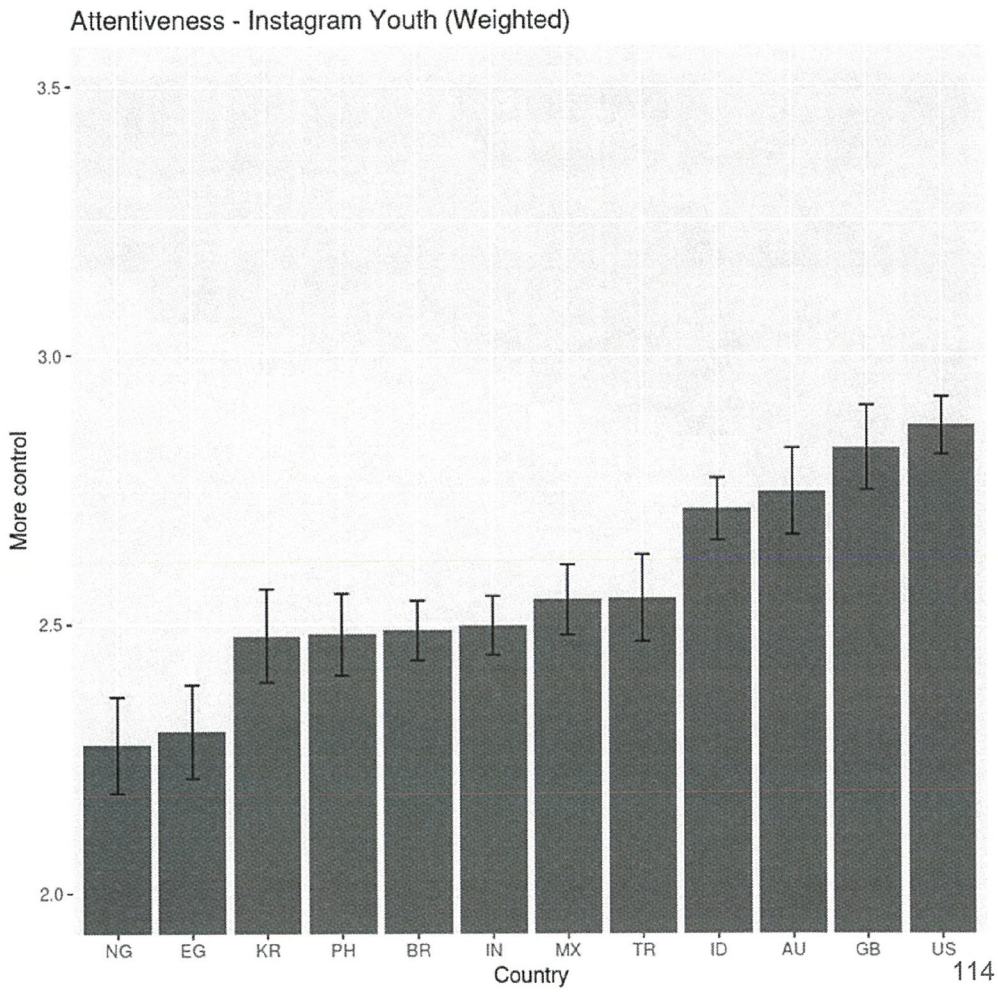


112

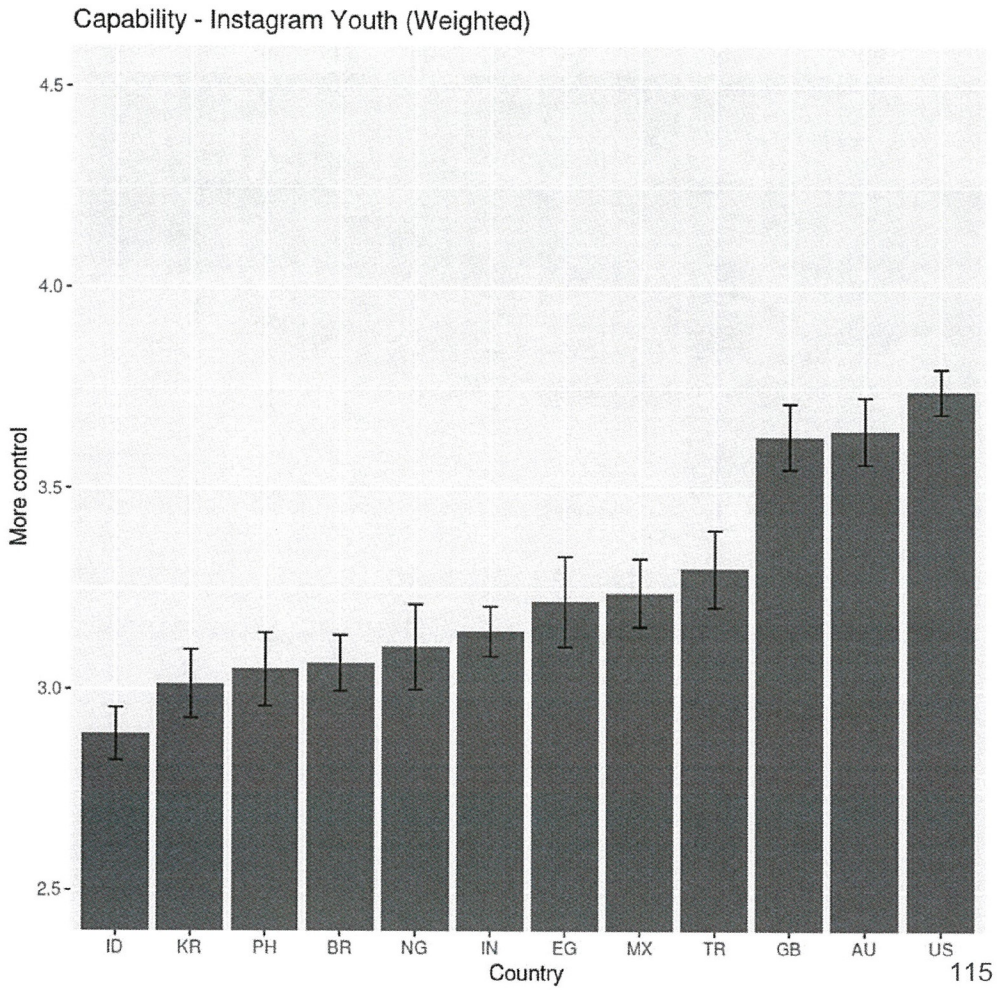
Facebook Adults: Mean perceived capability by country



Instagram Youth: Mean perceived attentiveness by country

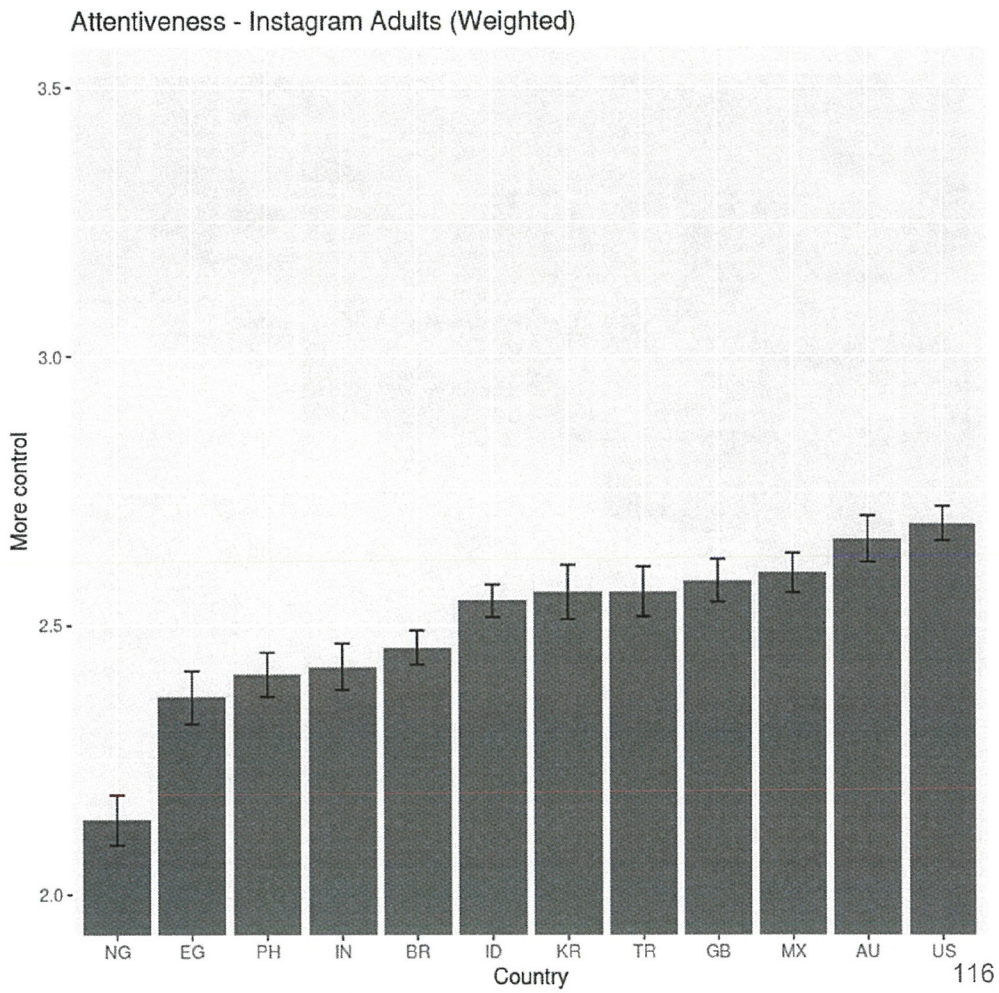


Instagram Youth: Mean perceived capability by country

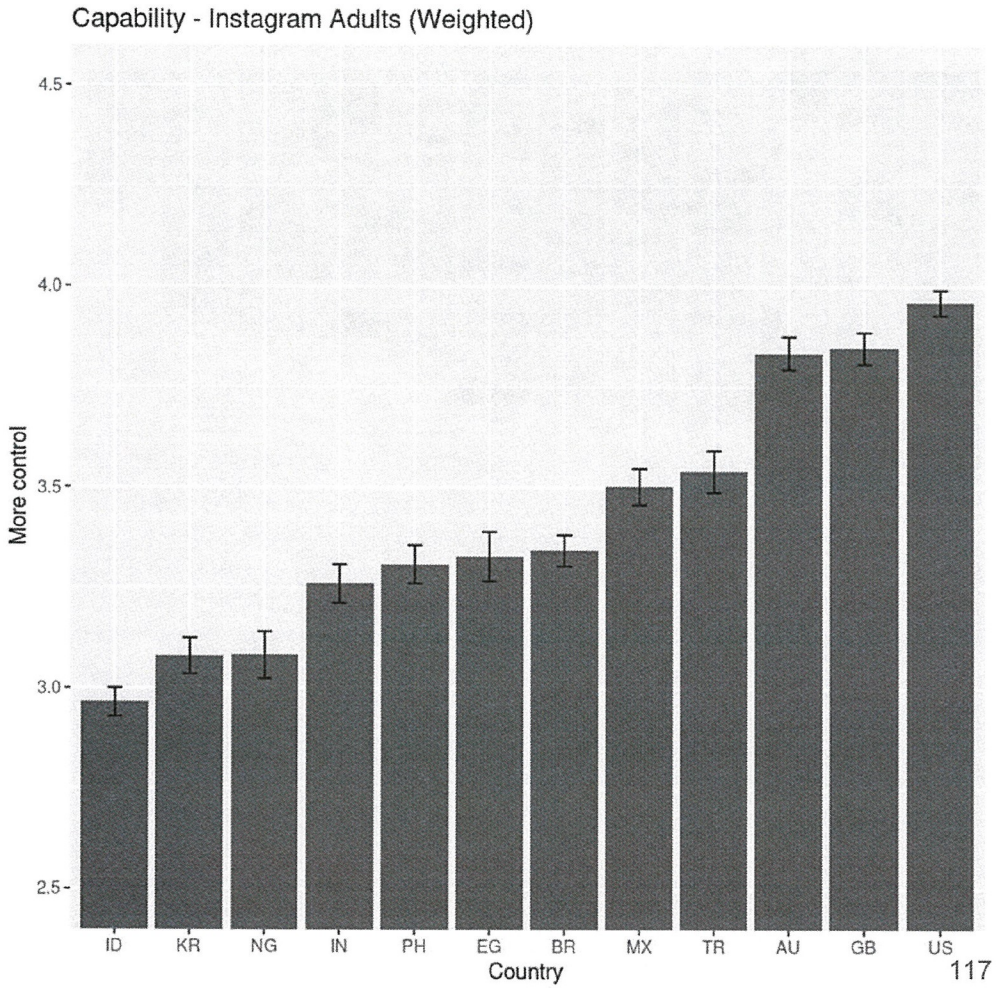


115

Instagram Adults: Mean perceived attentiveness by country



Instagram Adults: Mean perceived capability by country

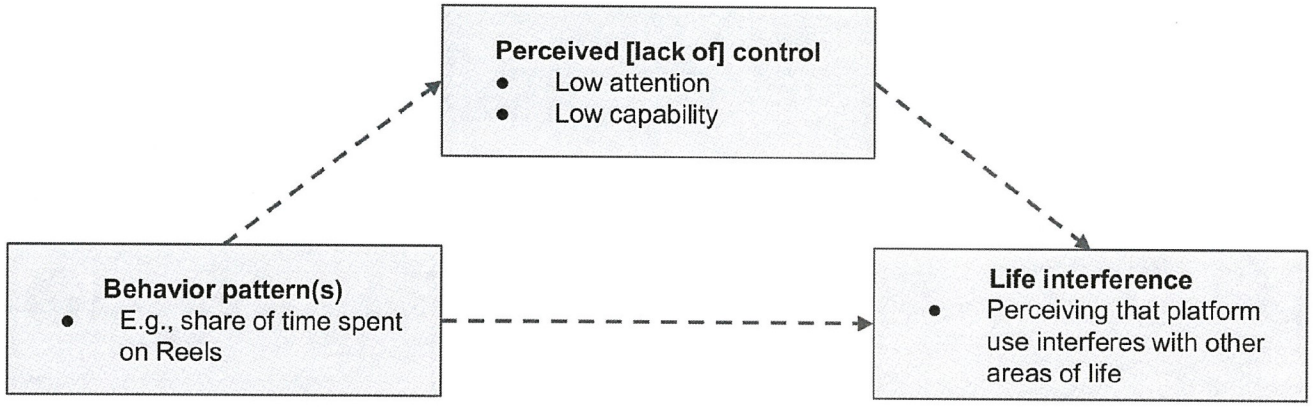


117

Older/extra slides

Mediation model

(note that dotted lines are **not** causal relationships; they are a statistical model of associations)



Problematic use

Definition

(note that attributing life impact to social media use is non-causal and subject to various biases)

Lack of control over social media use *perceived to contribute to* **Negative life impacts**

Next steps for analysis

Surfaces:

Specify details of product associations including strength of association and details of the behavior (data dictionary)

Life interference associations? + mediation model

Deep dive for the content variable

Deep dive for decision variable; any moderation?

Descriptives:

- Weighting
- Control for country; country differences

Heuristics:

Tom: wants to see the lasso regression and how well we can predict control scores from all behavioral signals (updated heuristics)

we can explain x amount of variance in control with y features

how well do these heuristics predict control (is a proxy possible?); what are updated operational metrics for product evals

FAQ: What is the broader set of success measures, beyond surveys?

Criteria	Inputs	IG MWB
Feature User Value Signals *The extent to which we have evidence that people who are exposed to the intervention in some way may derive value from it.*	Product awareness, adoption and retention. Perceived user value (survey or qualitative feedback) Behavioral signals of intervention impact	**P0** Specifically product engagement (adoption and retention) and behavior change associated with product usage
Ecosystem Metrics *The extent to which the intervention significantly moves key team/organizational engagement metrics.*	App, Surface (e.g. Feed), and Team (e.g. FI Controls) critical metrics	IG critical metrics
Well-being signals *The extent to which the intervention moves operationalized well-being- related measures in the right direction.*	Logged PU/Control behavioral correlates Perceived control survey	**P1** Our contribution: "Understand" inputs to guide roadmap prioritization and next investments (frequent checking behavior)
Legitimacy Value *The extent to which the intervention brings us at least to industry parity, and/or we have evidence to suggest that external experts support the work.*	Industry standards (audit) Expert recommendations	XMWB contribution: Audit, expert interviews, and background research provided scaffolding and product steer for multi-half strategy

See product ship decision framework[122]

Overall behavior patterns

27 ttention predictors on FB: comparison to past heuristic work

- Consistent indicators:
  - No engagement repeat sessions
  - Time spent + sessions
  - Push sessions (if overlapping with jewel mark as seen?)
- Overlapping indicators:
  - Surface switches (general; not specific to late night)
  - Passive sessions (general and evening; not specific to late night)
- Not replicated:
  - Short sessions
  - Jewel mark as seen
- New indicators:
  - Push sessions

Capability predictors on FB: comparison to past heuristic work

- Consistent indicators:
  - No engagement repeat sessions
- New indicators:
  - Push sessions

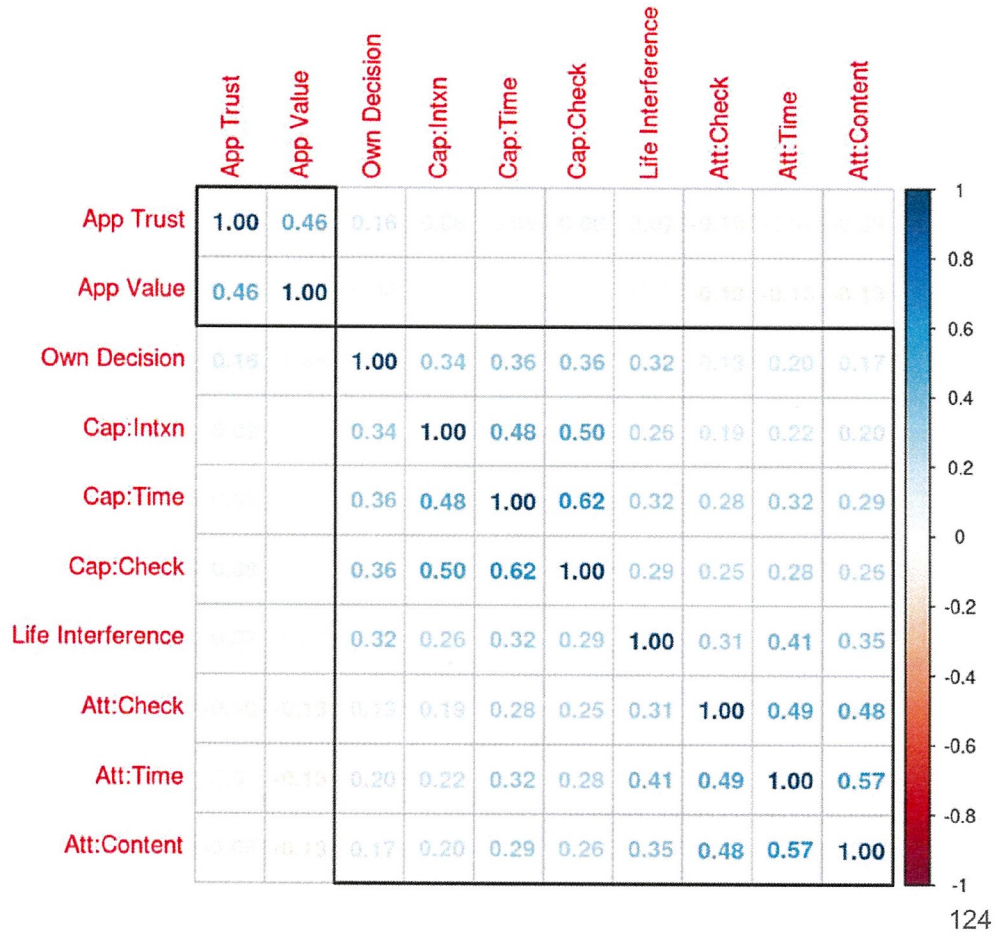
Slide 123

45 jewel mark as seen and push sessions- are these actually the same?
Elena Davis, 12/23/2022

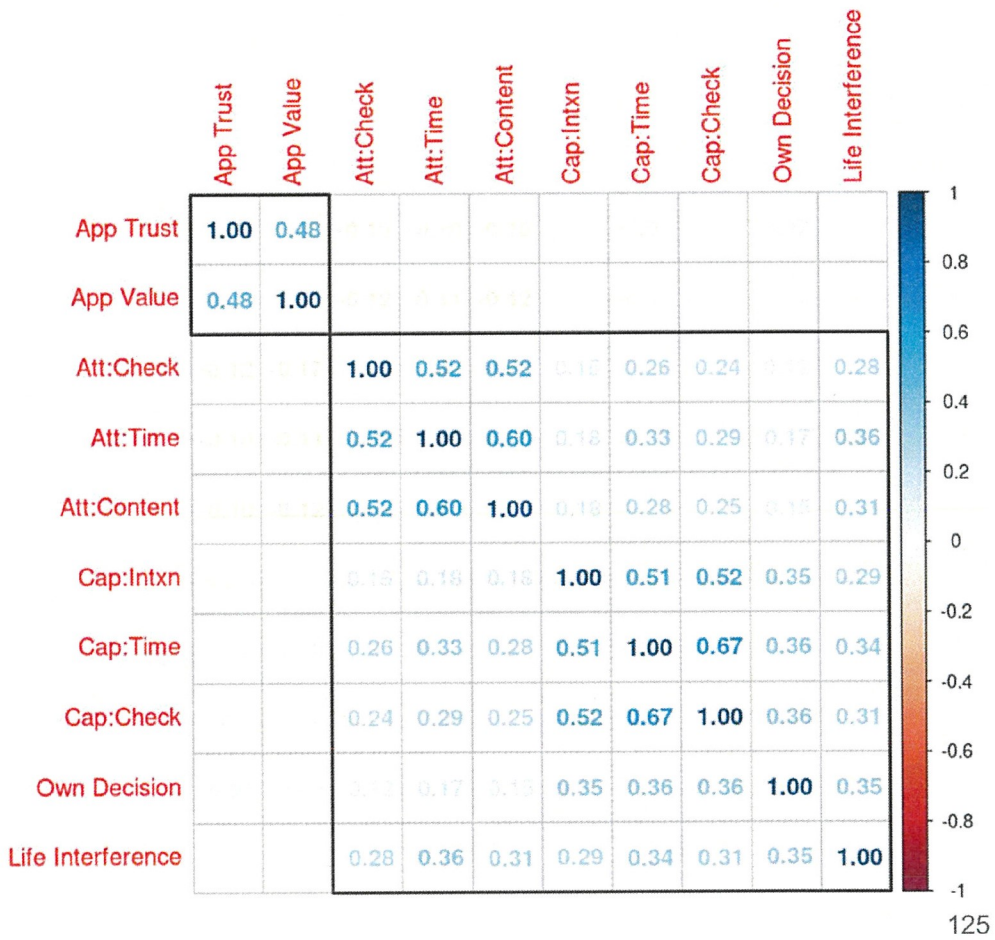
27 We have both of them. The notifications/jewel mark are captured by visits to the notif surface (they aren't identical but they are very correlated metrics.

push notifications are separate, they both originate from notifications but are not the same thing
Josh [redacted] 12/23/2022

Facebook Youth Correlations: All variables

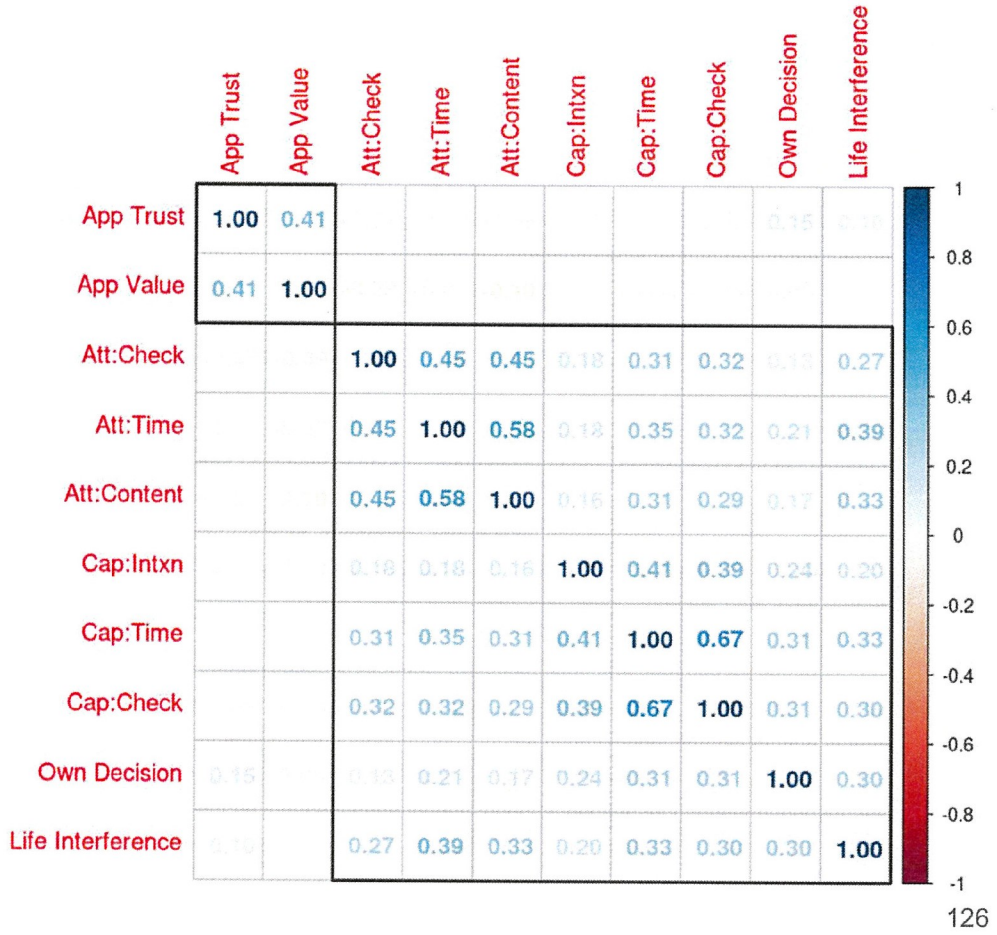


Facebook Adult Correlations: All variables



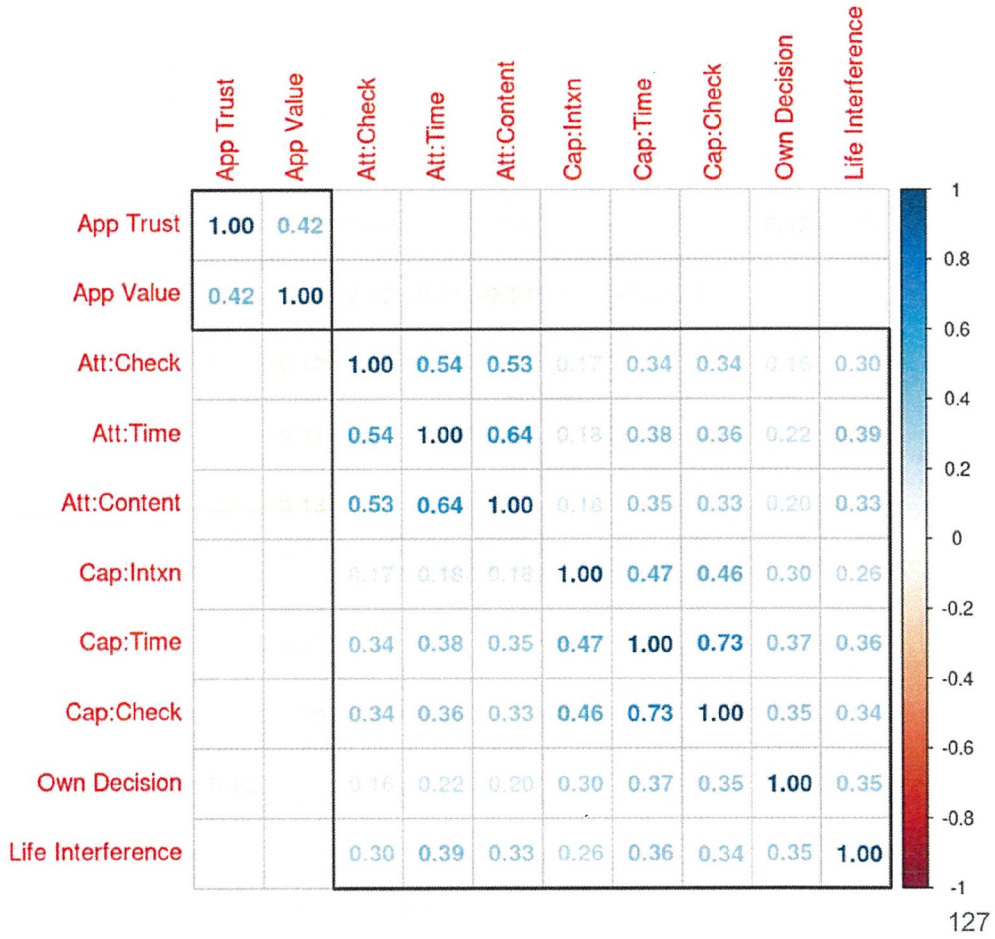
125

Instagram Youth Correlations: All variables



126

Instagram Adult Correlations: All variables



127

Descriptives - Facebook (*not weighted)

Higher scores = 'better' (more control or less life interference)

	Average 'attention' score *(composite of 3 items)*	% reporting low attention *(lowest control response to 1 or more attention items-> NOT WEIGHTED)*	Association with life interference	Average 'capability' score *(composite of 3 items)*	% reporting lowest capability *(lowest control response to 1 or more capability items-> NOT WEIGHTED)*	Association with life interference	% reporting high life interference *(highest frequency response of interference; 'all the time'--> NOT WEIGHTED)*
Facebook							
Adults N=30,609	2.60 (1.02)	33.8%	Adults with low attention report 20% higher life interference (4.1 vs. 3.3)	3.44 (1.09)	16.1%	Adults with low capability report 17% higher life interference (3.9 vs. 3.25)	7.5%
Youth N=29,361	2.71 (1.11)	38.2%	Youth with low attention report 23% higher life interference (3.9 vs. 3.0)	3.30 (1.16)	21.0%	Youth with low capability report 16% higher life interference (3.7 vs. 3.0)	11.9%
Comparison	t(46145)=-11.1 p<.0001 Youth reporting *higher* attention on average	Chisq(1)=99.4 p<.0001 Youth more likely to report *lowest* attention		t(38043)=12.2 p<.0001 Youth reporting *lower* capability on average	Chisq(1)=164.1, p<.0001 Youth more likely to report *lowest* capability		

Descriptives - Instagram (*not weighted)

Higher scores = 'better' (more control or less life interference)

	Average 'attention' score *(composite of 3 items)*	% reporting low attention *(lowest control response to 1 or more attention items-> NOT WEIGHTED)*	Association with life interference	Average 'capability' score *(composite of 3 items)*	% reporting lowest capability *(lowest control response to 1 or more capability items-> NOT WEIGHTED)*	Association with life interference	% reporting high life interference *(highest frequency response of interference; 'all the time'-> NOT WEIGHTED)*
Instagram							
Adults N=30,437	2.49 (0.98)	35.5%	Adults with low attention report 20% higher life interference (4.0 vs. 3.2)	3.43 (1.05)	18.3%	Adults with low capability report 17% higher life interference (3.9 vs. 3.2)	6.3%
Youth N=30,678	2.56 (1.00)	38.5%	Youth with low attention report 19% higher life interference (3.8 vs. 3.0)	3.27 (1.05)	14.8%	Youth with low capability report 15% higher life interference (3.6 vs. 3.0)	9.2%
Comparison	t(51815)=-7.68 p<.0001 Youth reporting *higher* attention on average	Chisq(1)=50.4 p<.0001 Youth more likely to report *lowest* attention		t(44326)=16.5, p<.0001 Youth reporting *lower* capability on average	Chisq(1)=97.0 p<.0001 Youth more likely to report *lowest* capability		

Descriptives takeaways - both FB and IG

- Perceptions of low attention are ~1.5-2x more common than perceptions of low capability
- Overall youth and adults report similar rates of low control (differing by a few percentage points); adults more likely to report low attention and youth more likely to report low capability
- Both low attention and low capability are associated with higher likelihood of reporting life interference from Facebook use (2-3x more likely); attention has larger associations

Approach: Updated survey measurement of potential problematic use

46bjective	Outcome
Identify a focus for problematic use that can be linked to on-platform behaviors (i.e., more likely to be associated with experiences occurring on-platform) Negative life impact as an outcome is more distal and subject to biases.	Focus on users' perception of control over their time and experiences on our platforms (note); explore how perceived control may be related perceiving that app use interferes with other responsibilities in life Begin assessments on Facebook and Instagram, where this topic has been prioritized to date (source)
Design and validate survey measures of users' perception of control, to ask users on-platform. Previous measurement (1, 2) had few questions about perceived control specifically, considered control as a unitary construct, and combined it with negative life impact.	Measures were developed to capture 2 types of low control over platform use (piloted + validated): 1) **Low attentiveness**: frequently spending more time, checking more, and consuming more content than intended, without realizing it 2) **Low capability**: perceived inability to reduce or limit time spent, checking, and interactions, even if wanting to

Slide 131

46 a bit of metric history re: why this type of measurement was needed (why not a single measure of well-being)
Elena Davis, 1/27/2023

Measurement to do (first for IG):

Total predictive accuracy if we throw everything at it

- To predict low attentiveness/ low capability (score of <=2); life interference (lowest 2 scores)
  - How well can we predict any of this? (put the proxy to bed)
  - (precision/recall overall)

Among the candidate heuristics in a simplified model (top 10?)

- Same exercise as above
- What might be a reasonable threshold for cut-off for goaling and/or targeting?
- Qs:
  - How sensitive is the unified predict that would be the best shot at predicting (upper bound of predictive accuracy)?
  - How sensitive to movement is each behavior?
    - (precision/recall)

How much of life interference is predicted by control; how much can we predict the REST of life interference form anything on platform

- Separate table of behaviors associated w life interference?

Future direction to think about prioritized metrics to focus on:

- SWOP metric – an index of behaviors stacked on top of each other (w deduplication – controlling for one of behaviors, how much does another set of behaviors 'matter' for predicting control?)
  - What is the right 'in-road' to changing behavior?

	Precision	Recall	AUC
Model with all features	0.17	0.73	0.66
Proposed PU heuristic	0.16	0.67	0.63

On-platform measurement of problematic use

Objective	Outcome
Identify logged behaviors associated with low control (low attentiveness and/or low capability), for teens and adults, across surfaces and products on Facebook and Instagram	Results in this deck from: N=30,699 adults and 30,346 youth on FB N=30,810 adults and 30,756 youth on IG Across 12 countries: AU, BR, EG, GB, KR*, ID, IN, MX, NG, PH, TR, US (*IG only)
Identify global behaviors associated with low control	Results in this deck on updated behavioral heuristics used for targeting and/or goaling PU interventions

Outline

Approach:

- Overall PU approach
- Measurement approach

Facebook results

Instagram results

Methods detail

Results detail

Outline

- Approach:
  - Measuring control
- Findings for teams:
  - Facebook section
    - Rates/importance (descriptives)
    - Associations with on-platform behaviors
  - Instagram section
    - Rates/importance (descriptives)
    - Associations with on-platform behaviors
- Measurement:
  - Why not proxy?
    - How little we can predict
  - What to use for measurement
    - Operational metrics/heuristics
      - Which are the top predictors? (agnostic of if we should use these for goaling?)
      - Description of next steps (measurement roadmap)

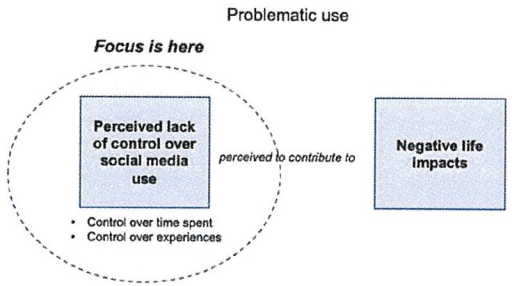
Summary of users' report of perceived low control

How common are experiences of perceived low control on Instagram?	Rates of reporting the *lowest* scores (<=2 on 5 pt composite scale; or 1 or more lowest responses on any item): **Low attentiveness:** ~34-40% of users overall **Low capability**: ~13-19% of users overall
How do these experiences differ across ages and markets?	Adults slightly more likely to report low attentiveness; youth slightly more likely to report low capability. (Differences of a few % points.) Markets differ: higher control in western/high income (US, GB, AU); lower in some emerging markets (NG, EG, ID) and also relatively low in KR - possibly due to cross-cultural differences in perceptions of control and role of Facebook in people's lives
How does perceived low control relate to other perceptions of the app?	As perceived low control increases, odds of reporting that their use of Instagram prevents the fulfillment of other responsibilities (i.e., 'life interference') are 2.7-3.0x for attentiveness, 1.9-2.0x for capability Low control is *not* strongly related to perceiving value from the app or views on whether the app is trustworthy or not (correlations < .14)

Descriptive details starting here

49

Platform behaviors are associated with perceiving 'life interference' primarily through perceived control



Attributing life impact to social media use is non-causal and subject to various biases. This means there are likely many other contributing factors unrelated to platform use.

This survey asked "How often does your use of Instagram prevent you from fulfilling your responsibilities in other areas of your life (for example, work, school, or relationships)?" - the MTS 'life interference' question.

With all on-platform signals, we can predict 17% of the variance in perceived 'life interference' for youth and 20% of the variance for adults.

The limited association between platform use (behavioral metrics) and perceived 'life interference' is almost entirely explained through low control for teens; and mostly explained for adults.

Beyond the effects of perceived low control, platform signals only explain 0.7% of the variance in teen users reported life interference and 1.8% of the variance for adults. This means for teens 96% of the relationship between the platform and perceived life interference is explained by perceived lack of control while for adults 91% of the relationship is explained by low control.

Slide 137

9 I don't follow this slide at all. This is also kind of random in the midst of talking about the model results. Can you slide this down to a later section or just put it in the appendix?
Judy [redacted] 2/6/2023

48 @[redacted]@meta.com to follow up; needs work
Elena Davis, 2/7/2023

28 This is addressing very specific push we got from IG DS and to some extent Research leads. They are not convinced about low perceived control whe life interference is most associated with how external stakeholders think about PU (idk if that framing is correct just the resistant we have faced).

This is meant to show
1. Life interference is related to our platform through low control. We don't see a relationship between behaviors/experiences on the platform relatec to life-interference other than through low control.

2. Low control is more proximal to the platform and has less noise than life interference so we can feel better about the strength of relationships with low control (and hopefully the heuristics will be better because of it).

This maybe isn't super clear here but the AUC from life interference prediction is 0.67 in Carlos prior work and 0.7 in this iteration compared with ~0.8 for low control.
Josh [redacted] 2/7/2023

10 I think if you wrote 1 and 2 just like you said here it would be clearer.
Judy [redacted] 2/7/2023

47 noting this heading needs work
Elena Davis, 2/7/2023

49 yayy also i was going to pull up the theory of change too. this is so helpful
Elena Davis, 2/7/2023

Takeaways

Overall lower control	Many behavioral predictors are associated with both forms of perceived low control, though predictive power for perceived low attentiveness is overall much higher than for perceived low capability. Top predictors track with high use (total time spent + numbers of sessions), additionally use in the evening and into late night. Predictors largely the same for adult and youth. → *Findings are consistent with the stance around focusing on late night use as important to support user control; extending to evening use may be a next step. Expanding solutions to adult cohorts may also be beneficial.*
Low attentiveness	Uniquely associated with video chaining; one hypothesis is that autoplaying of subsequent videos could be related to behavior done 'without realizing it' → *There may be specific attentiveness nudges or autoplay settings relevant to this aspect of control on video surfaces.*
Low capability	Uniquely associated with short sessions; one hypothesis is this could be related to checking behavior. → …

Descriptives

How common are experiences of perceived low control on Facebook?	Rates of reporting the *lowest* scores (<=2 on 5 pt composite scale; or 1 or more lowest responses on any item): **Low attentiveness:** ~30-34% of users overall **Low capability**: ~16-20% of users overall
How do these experiences differ across ages and markets?	Adults slightly more likely to report low attentiveness; youth slightly more likely to report low capability. (Differences of a few % points.) Markets differ: higher control in western/high income (US, GB, AU); lower in emerging markets (PH, NG, EG) - possibly due to cross-cultural differences in perceptions of control and role of Facebook in people's lives
How does perceived low control relate to other perceptions of the app?	As perceived low control increases, odds of reporting that their use of Facebook prevents the fulfillment of other responsibilities (i.e., 'life interference') are 2.4-2.7x for attentiveness, 1.7-1.8x for capability Low control is *not* strongly related to perceiving value from the app or views on whether the app is trustworthy or not (correlations < .15)

DOCUMENT SLIPSHEET

Bates Number: META3047MDL-074-00027496

Custodian: DAVIS_ELENA; [redacted]_ANYA; [redacted]_BRIAN

Author: [redacted]@META.COM

Filename: -A-C PRIV--SENSITIVE -DO NOT SHARE- PU CONCENTRATION SURVE_1J9BU5MPVGXO4X3KQN0WWLD5NZD0VK8JB7G_4MUJAEOM.PPTX

Create Date: 12/1/2022 12:00 AM

Last Modified Date: 2/14/2024 12:00 AM

TNAG BEGBATES:

TNAG ENDBATES:

MAAG BEGBATES:

MAAG ENDBATES:

File Path: /[redacted]_ANYA; /DAVIS_ELENA; /[redacted]_BRIAN

Confidentiality Designation: HIGHLY CONFIDENTIAL (COMPETITOR)

META_Pointee:

META_Pointer: