GOAL RESEARCH

SUMMARY

In this paper, I am exploring our daily diary data. I am looking at:

- Descriptive Statistics
    - Which types of goals are most frequent among the study population
    - Which types of goals have highest completion, effort, affect etc.
- How goal type is associated with grades
- Goal Hierarchy:
    - How similarity between daily and monthly goals vary as the month progresses
- How affect is related to goal goal completion
- How similarity between today and tomorrow's goals are associated with goal completion tomorrow
    - How today's completion influences this relationship
    - How effort is related to goal completion
    - How today's affect influences this relationship

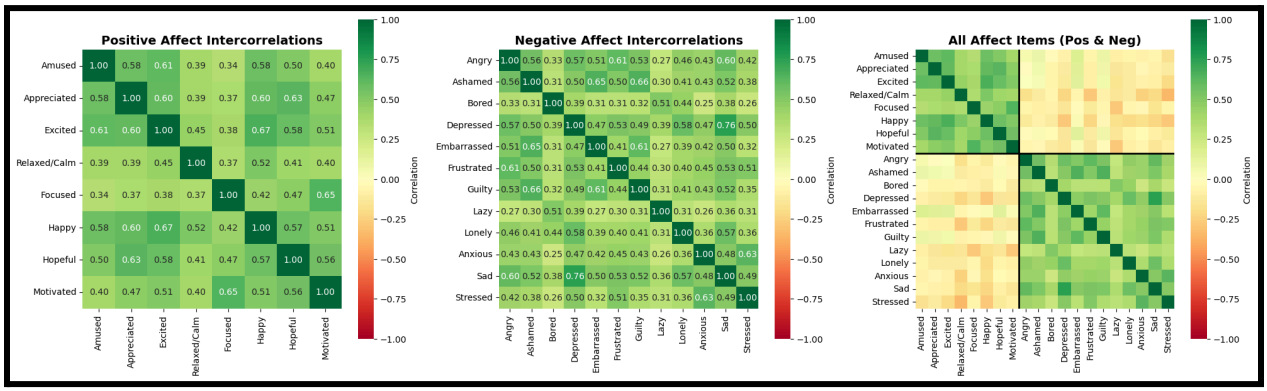
DATA

We had Daily diary measurements for ~183 participants | ~9000 days (across two cohorts | 3 months each)

- **Preprocessing**:
    - Goal Alignment:
        - Participants entered two goals each evening (Day t) for the following day (Day t + 1). To align goals with the corresponding behavioral and affective outcomes, I shifted all goal timestamps forward by one day. Thus, each record reflects the day when the goals were executed and the associated completion and emotion measures were logged.
    - Goal Count: ensured at least 40 days worth of entry for each participant retained.
    - Final dataset had **112 participants with ~7800 total days**

- **Categorizing Goals:** Each of the two goals per day was categorized into 1 of 35 categories by OpenAI GPT 4
- **Binary Goal Similarity:** Binary similarity between today and tomorrow's goals (4 values) was generated using Open AI GPT 4
- **Similarity:** Continuous similarity between today and tomorrow's goals (4 values) was generated via cosine similarity on the goal text embedding from Open AI.
- **Affect measures**:
    - Positive affect: 8 items (amused, appreciated, excited, relaxed/calm, focused, happy, hopeful, motivated)
    - Negative affect: 12 items (angry, ashamed, bored, depressed, embarrassed, frustrated, guilty, lazy, lonely/isolated, nervous/anxious, sad, stressed)

- Correlation analysis showed positive items strongly correlated with each other (r = 0.34-0.67), negative items strongly correlated with each other (r = 0.25-0.76), and near-zero correlation between positive and negative items (mean r = -0.05)



- Internal consistency (Cronbach's Alpha): Pos affect α = 0.89; Neg affect α = 0.90
- Created composite scores by averaging items within each scale

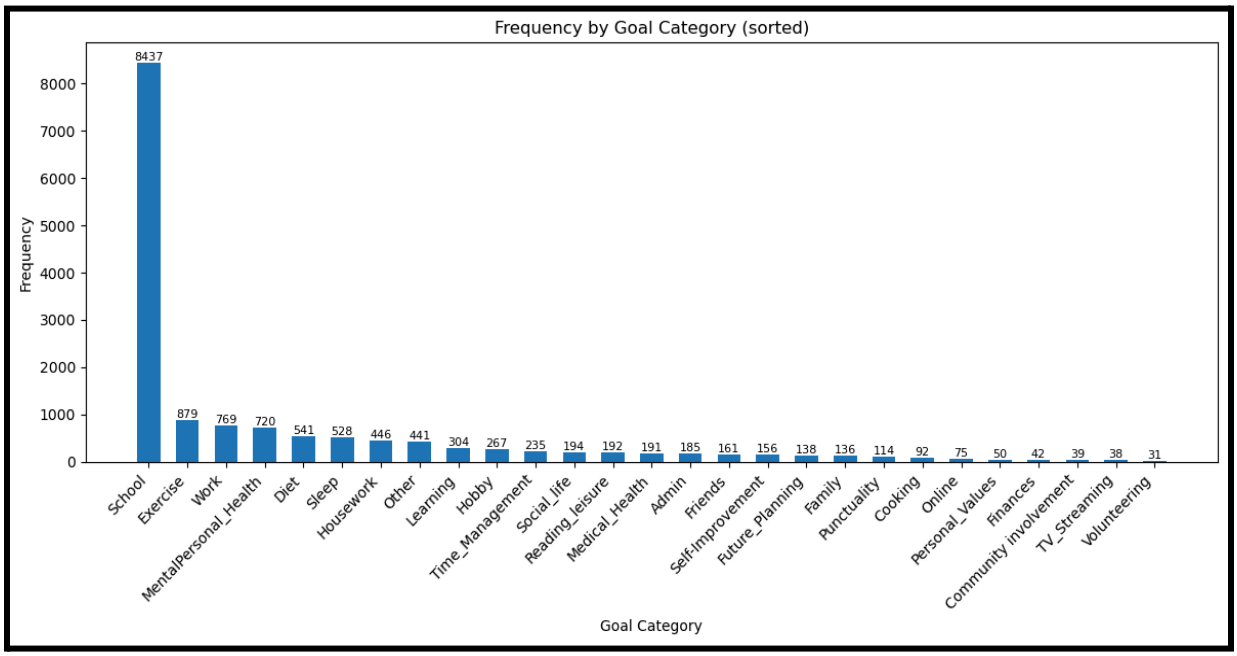
EXPLORATIONS

1. Descriptive Statistics
2. Goal Type x Grades
3. How similarity between daily and monthly goals vary as the month progresses
4. How affect is related to goal goal completion
5. Similarity (today & tomorrow) x Completion Today -> Completion Tomorrow

1. Descriptive Statistics

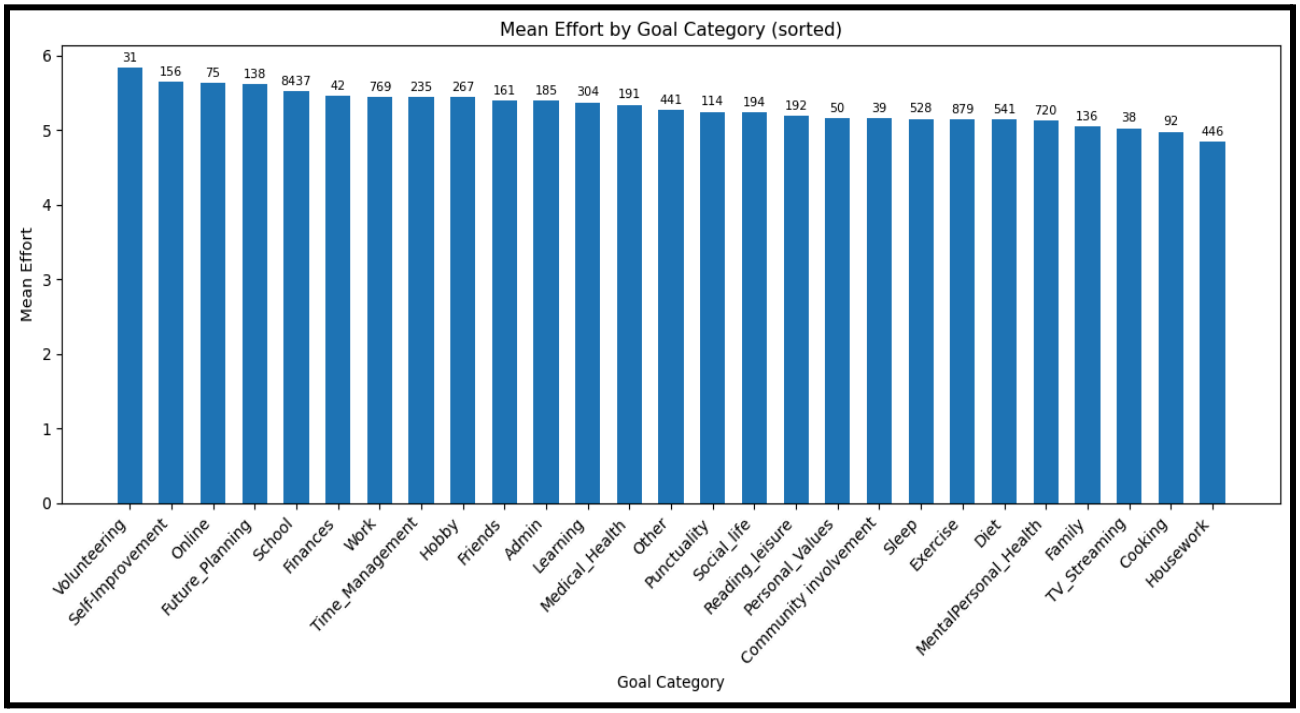
FREQUENCY OF EACH GOAL CATEGORY

	label	count	effort	completion
0	School	8437	5.522698	60.479080
1	Exercise	879	5.142207	64.745165
2	Work	769	5.440832	69.453836
3	MentalPersonal_Health	720	5.129167	67.002778
4	Diet	541	5.140481	66.064695
5	Sleep	528	5.149621	69.229167
6	Housework	446	4.840807	60.376682
7	Other	441	5.263039	67.113379
8	Learning	304	5.371711	45.832237
9	Hobby	267	5.438202	61.378277
10	Time_Management	235	5.438298	58.204255
11	Social_life	194	5.237113	77.216495
12	Reading_leisure	192	5.187500	59.635417
13	Medical_Health	191	5.340314	73.560209
14	Admin	185	5.389189	55.362162
15	Friends	161	5.397516	77.136646
16	Self-Improvement	156	5.641026	58.615385
17	Future_Planning	138	5.615942	51.920290
18	Family	136	5.051471	73.602941
19	Punctuality	114	5.245614	69.026316
20	Cooking	92	4.978261	60.576087
21	Online	75	5.626667	51.680000
22	Personal_Values	50	5.160000	76.500000
23	Finances	42	5.452381	55.214286
24	Community involvement	39	5.153846	66.461538
25	TV_Streaming	38	5.026316	72.000000
26	Volunteering	31	5.838710	66.548387

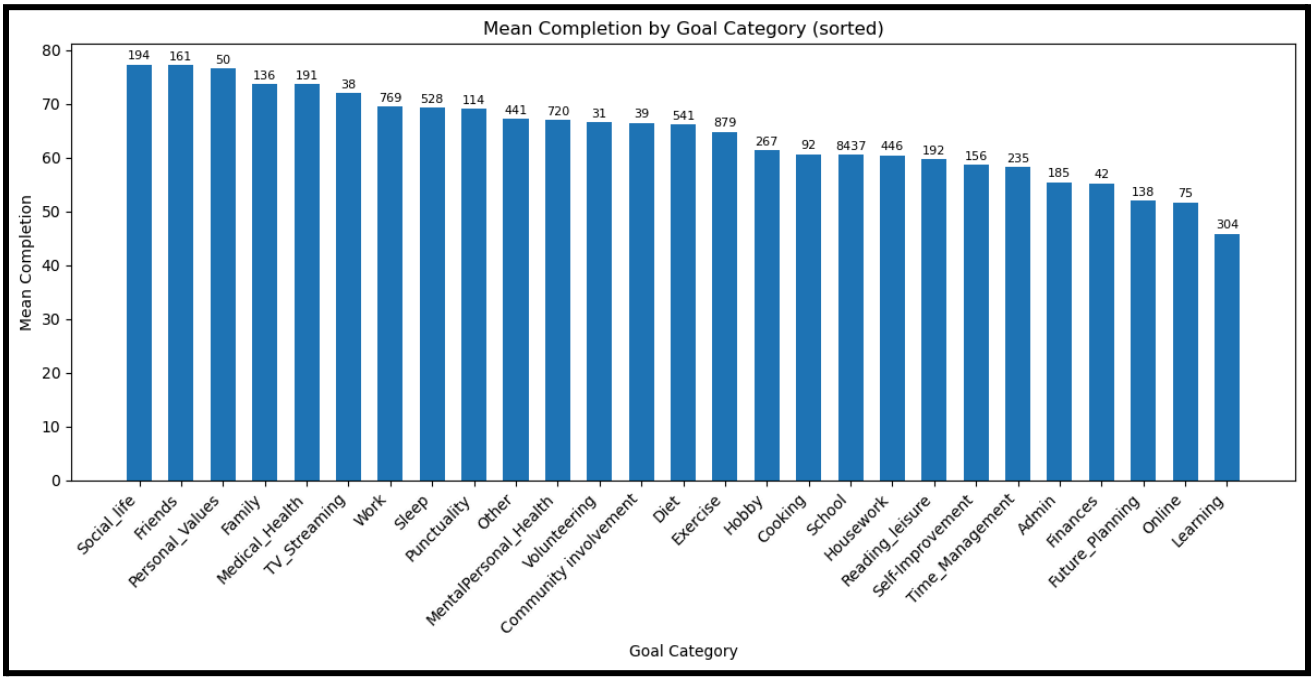


MEAN EFFORT PER GOAL CATEGORY

(Count given on top of each bar)



MEAN COMPLETION PER GOAL CATEGORY

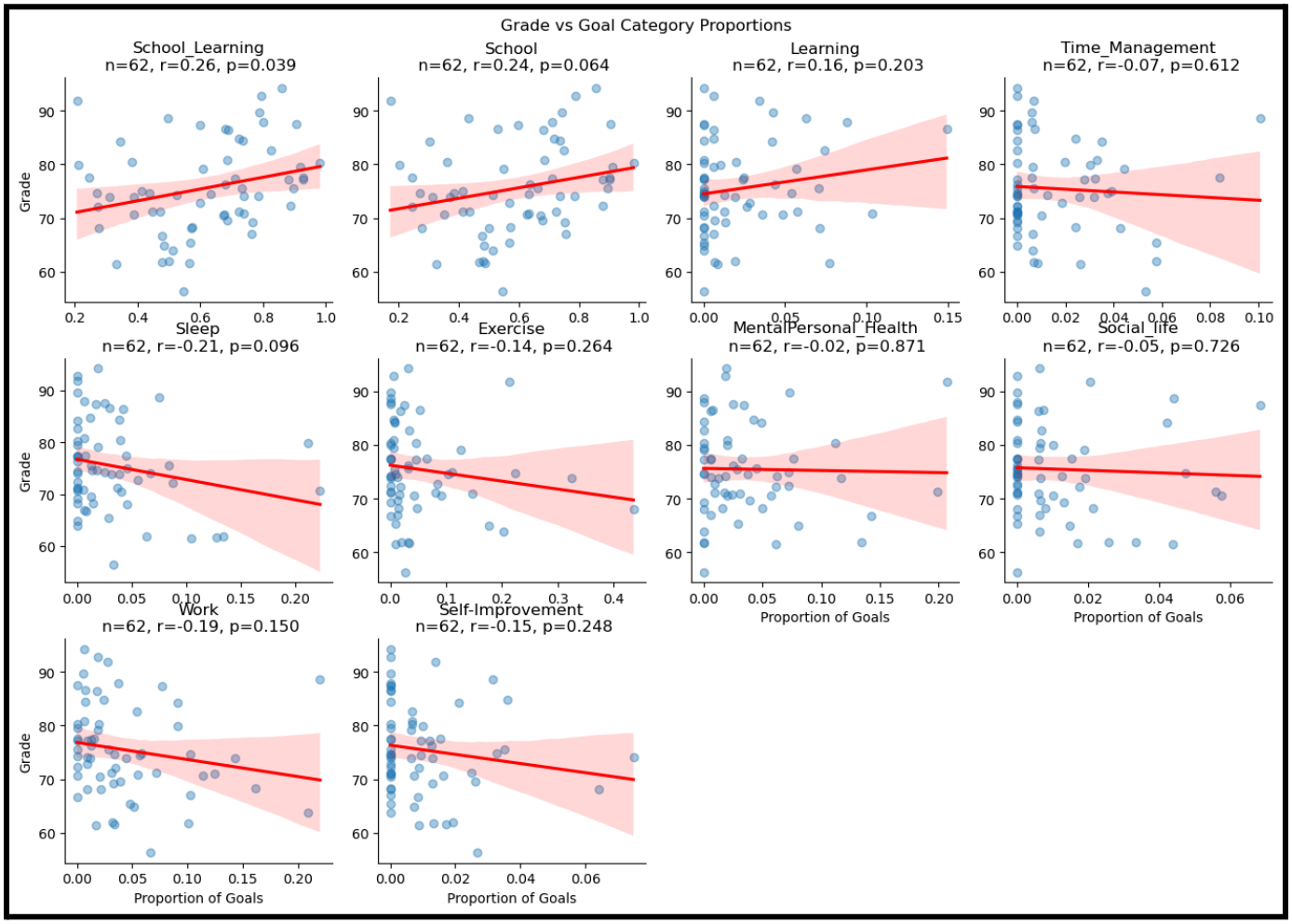


2. Goal Type x Grades

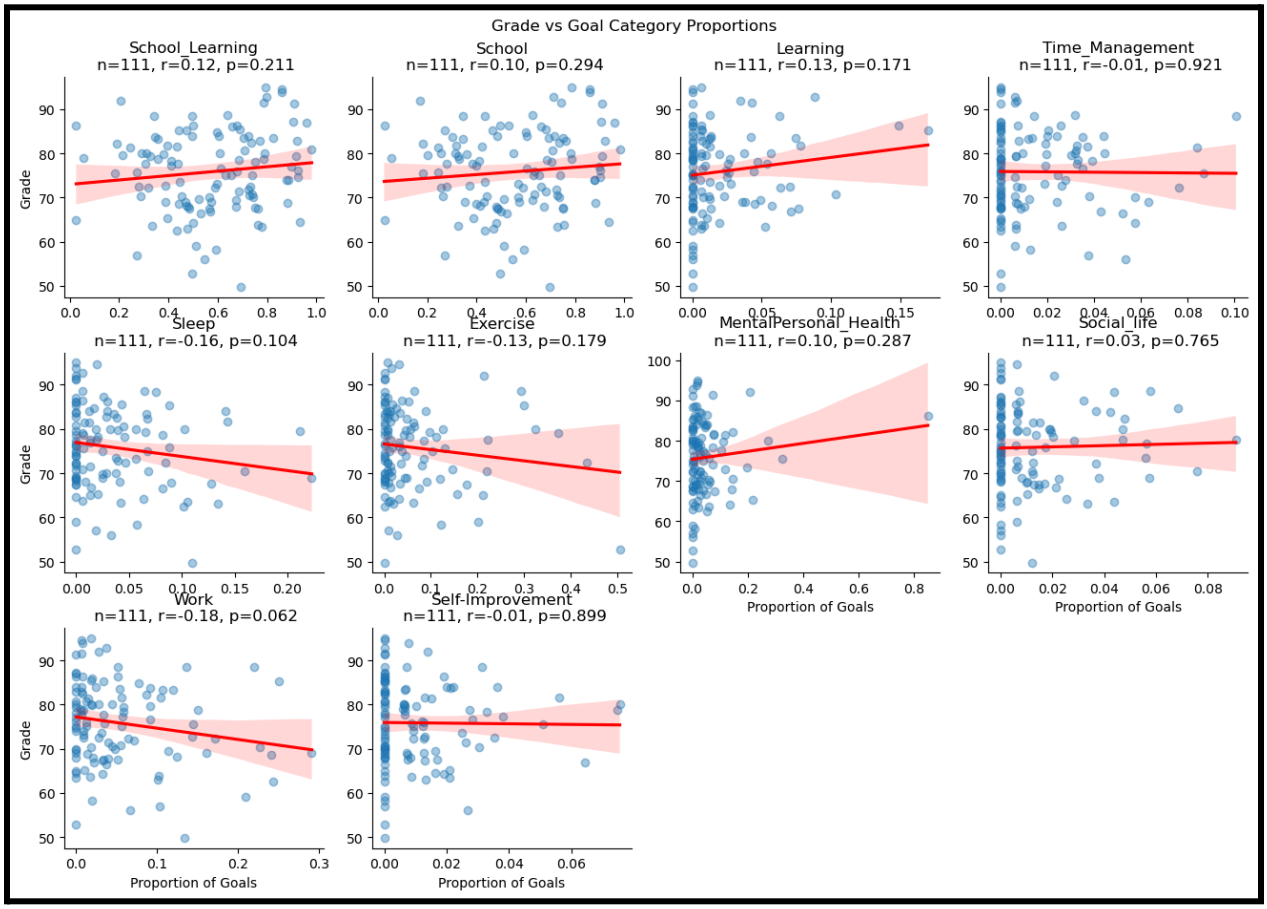
The aim is to explore whether the proportion of different goal types across the semester is related to students' final grades.

For example, Participant A had 56% of their daily goals related to school and earned a final grade of 98%. Participant B had 40% school-related goals and a grade of 77%. We plotted the proportion of school goals against final grades for each participant, and repeated this analysis for other goal types such as work, sleep, and exercise.

ONLY COHORT 1



BOTH COHORTS



3. Goal Hierarchy

AREA 3: INDIVIDUAL DAILY GOAL SIMILARITY WITH MONTHLY GOALS

Columns

- **ParticipantIdentifier**
- **trial_date, monthly_goal_set_date**
- **month_number**
- **daily_goal_1, daily_goal_2**
- **MONTHLY_goal_1, MONTHLY_goal_2**
- **DAILY_goal1_label, DAILY_goal2_label**
- **daily_monthly_sim: max between the four similarity scores**
- **chosen_daily: chosen daily goal for max similarity**
- **chosen_monthly: chosen monthly goal for max similarity**

I fit linear mixed-effects models to examine the relationship between

- Day_since_month_start
- Daily_monthly_sim

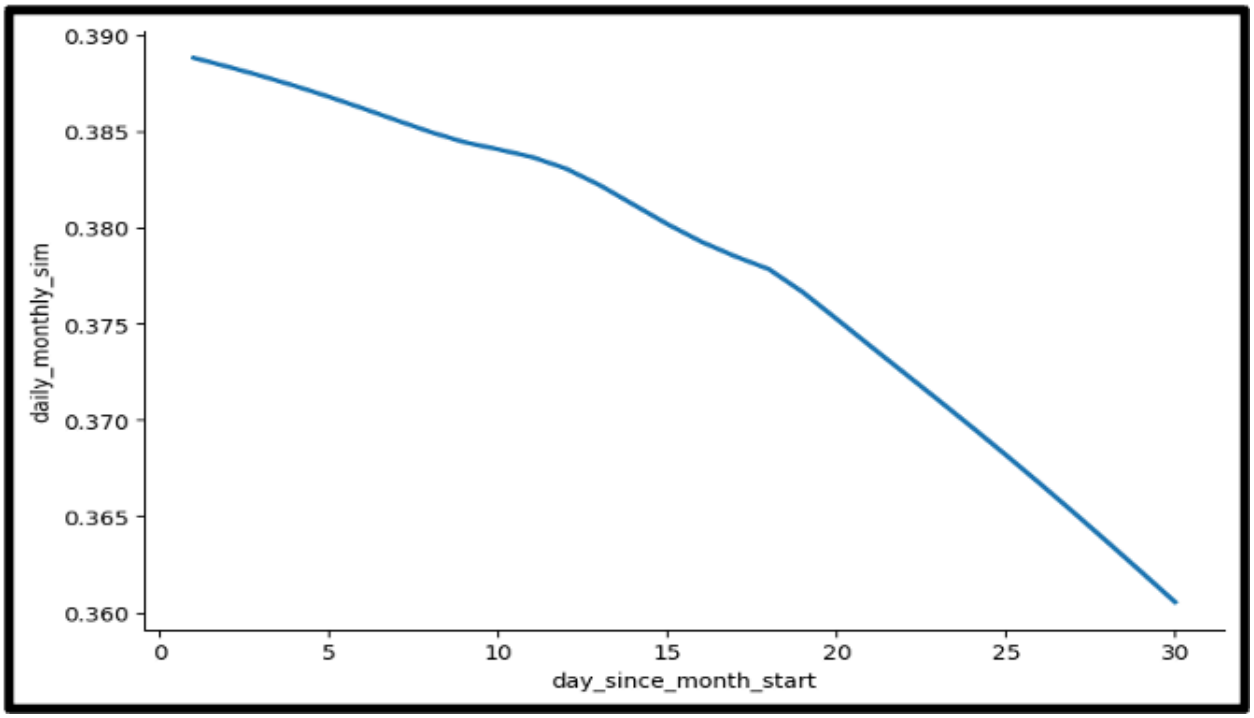
with participants included as a random intercept to account for repeated measurements.

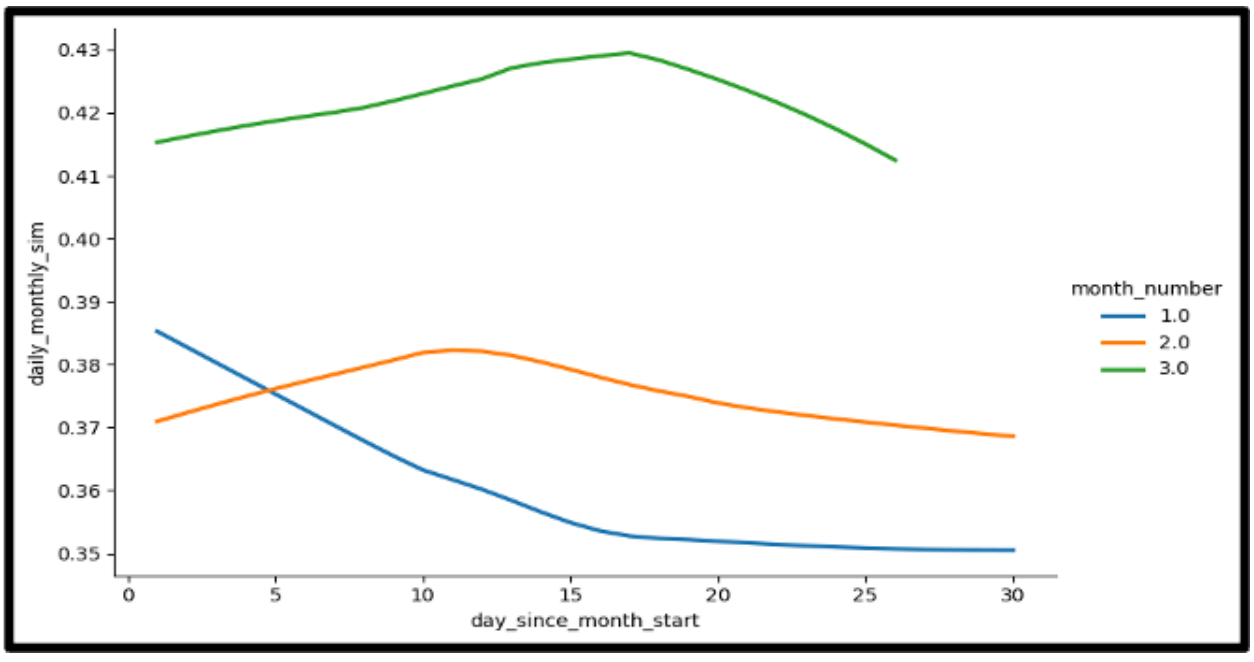
The analysis was conducted under two filtering criteria:

- Participants with at least 40 total entries spanning three or more months
- Participants with at least 15 entries per month

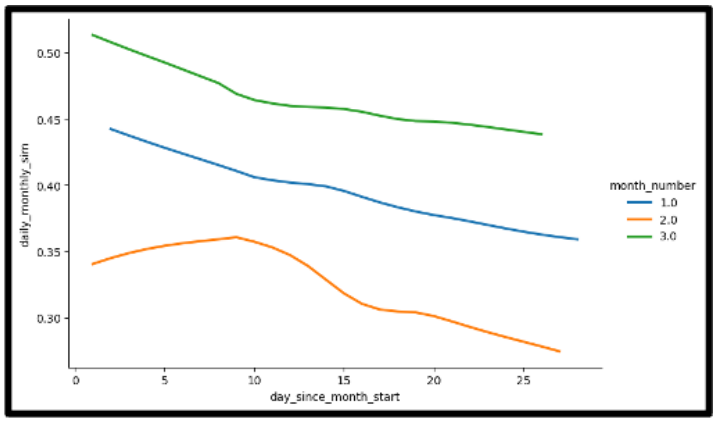
Across both criteria, there was a significant negative association between *day_since_month_start* and *daily_monthly_sim*, indicating that daily goals became progressively less aligned with monthly goals as the month progressed. However, this effect was driven primarily by the first month: the association was significant in Month 1 but attenuated and not statistically significant in Months 2 and 3.

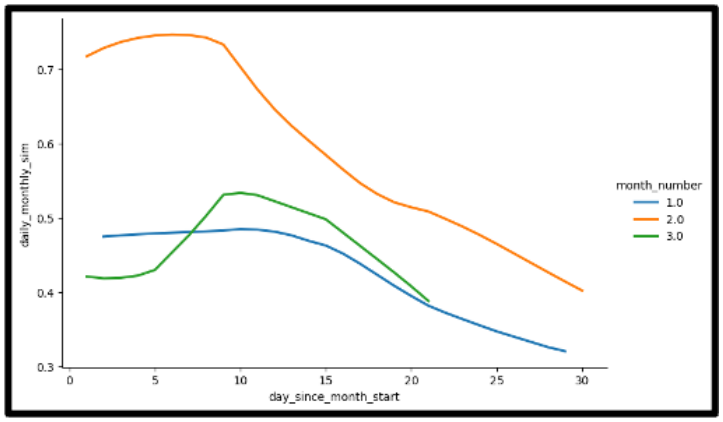
On average, across all months and people, what does similarity look like as a function of day since the start of a month?

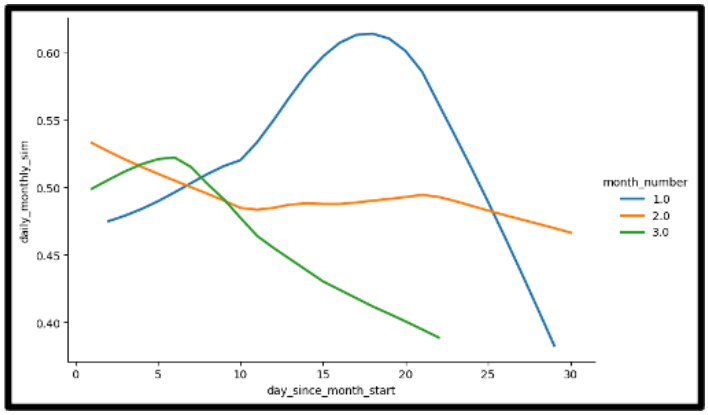


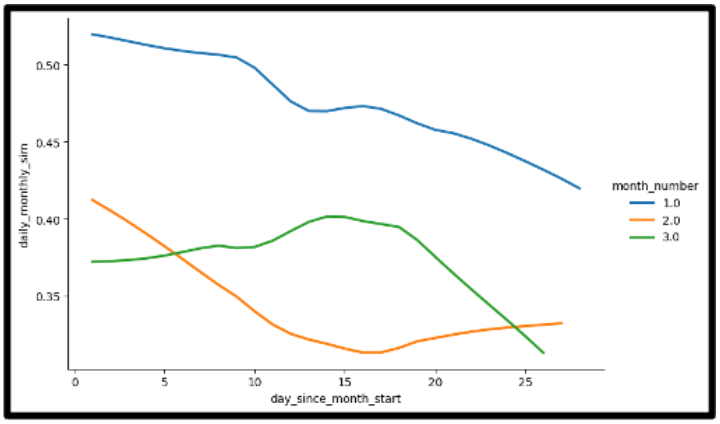


Let's plot some participant specific plots.









Mixed Linear Model Regression Results

- Participants with at least 40 total entries spanning three or more months

Model:	MixedLM	Dependent Variable:	daily_monthly_sim			
No. Observations:	7170	Method:	REML			
No. Groups:	103	Scale:	0.0235			
Min. group size:	40	Log-Likelihood:	3082.9637			
Max. group size:	81	Converged:	Yes			
Mean group size:	69.6					

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	0.441	0.012	38.260	0.000	0.418	0.463
day_since_month_start	-0.001	0.000	-5.747	0.000	-0.002	-0.001
Group Var	0.012	0.012				

Mixed Linear Model Regression Results

- Participants with at least 15 entries per month

Model:	MixedLM	Dependent Variable:	daily_monthly_sim			
No. Observations:	5714	Method:	REML			
No. Groups:	76	Scale:	0.0247			
Min. group size:	58	Log-Likelihood:	2307.8315			
Max. group size:	81	Converged:	Yes			
Mean group size:	75.2					

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	0.454	0.015	30.910	0.000	0.425	0.483
day_since_month_start	-0.001	0.000	-4.708	0.000	-0.002	-0.001
Group Var	0.015	0.016				

4. Affect ~ Completion

Model 1: tomorrow_completion ~ tomorrow_pos_affect

Table 1: Affect Predicting Goal Completion

Predictor	Model 1: Positive Affect	Model 2: Negative Affect
Fixed Effects		
Intercept	43.05*** (1.82)	71.64*** (1.74)
Tomorrow's Positive Affect	7.29*** (0.51)	—
Tomorrow's Negative Affect	—	-4.19*** (0.54)
Random Effects		
Participant Variance	157.83	168.57
Residual Variance	1207.63	1219.22
Model Fit		
Log-Likelihood	-72708.26	-72780.92
N observations	14,606	14,606
N groups	112	112

- Higher positive affect today is related to higher goal completion today
- Relationship reverses for negative affect

Model 2: tomorrow_pos_affect ~ today_completion × similarity

- For all data, when similarity between today and tomorrow is high, then higher goal completion today leads to higher positive affect for tomorrow.
- No significant relation when data is separated into continuation and non continuation goals.
- Seems like noise
- Simpsons paradox

5. Coasting / Momentum

We have a Dataframe somewhat like this

PID	TODAY DATE	TOMORROW GOAL	TOMOROW GOAL ID	TODAY GOAL 1	TODAY GOAL 2	MAX SIM GOAL 1	MAX SIM	MAX SIM TODAY COMPLETION	TOMORROW COMPLETION
199	Aug 1	A	GOAL 1	X	Y	FALSE	0.80 (A-Y)	80 (Y)	40 (A)
199	Aug 1	B	GOAL 2	X	Y	TRUE	0.90 (B-X)	50 (X)	35 (B)
199	Aug 2			A	B				

Questions:

- Does high completion on a goal today lead to higher completion on a similar goal tomorrow (momentum), or does it lead to lower completion (compensation)

- Are multi-day spanning goals confounding the results? Do relevant effects hold after removing continuation goals? Does affect mediate any of these effects?

Models we are testing here

- **Model 1**: $\text{completion}_{t+1} \sim \text{similarity\_binary}_{t \times [t+1]} \times \text{completion}_{t}$
- **Model 2**: $\text{completion}_{t+1} \sim \text{similarity\_cont}_{t \times [t+1]} \times \text{completion}_{t}$
- **Model 3**: $\text{similarity\_cont}_{t \times [t+1]} \sim \text{effort}_{t} \times \text{completion}_{t}$
- **Model 4:** $\text{completion}_{t+1} \sim \text{similarity\_cont}_{t \times [t+1]} \times \text{completion}_{t} + \text{pos\_affect}_{t+1}$

Table 1: Main Models Predicting Tomorrow's Goal Completion (All Data)

- Model 1: **completion $_{t+1}$ ~ similarity_binary $_{t\ x\ [t+1]}$ x completion $_{t}$**
- Model 2: **completion $_{t+1}$ ~ similarity_cont $_{t\ x\ [t+1]}$ x completion $_{t}$**

Predictor	Model 1: Binary Similarity (Open AI)	Model 2: Cont Similarity
FIXED EFFECTS		
Intercept	59.17*** (1.37)	64.01*** (1.63)
Today's Completion	0.074*** (0.01)	-0.08*** (0.02)
Goal Similarity (Binary)	-10.08*** (1.12)	—
Goal Similarity (Continuous)	—	-17.12*** (1.91)
TWO WAY INTERACTIONS		
Today's Completion × Similarity	0.24*** (0.02)	0.40*** (0.03)
MODEL FIT		
N observations	14,660	14,660
N groups	112	112

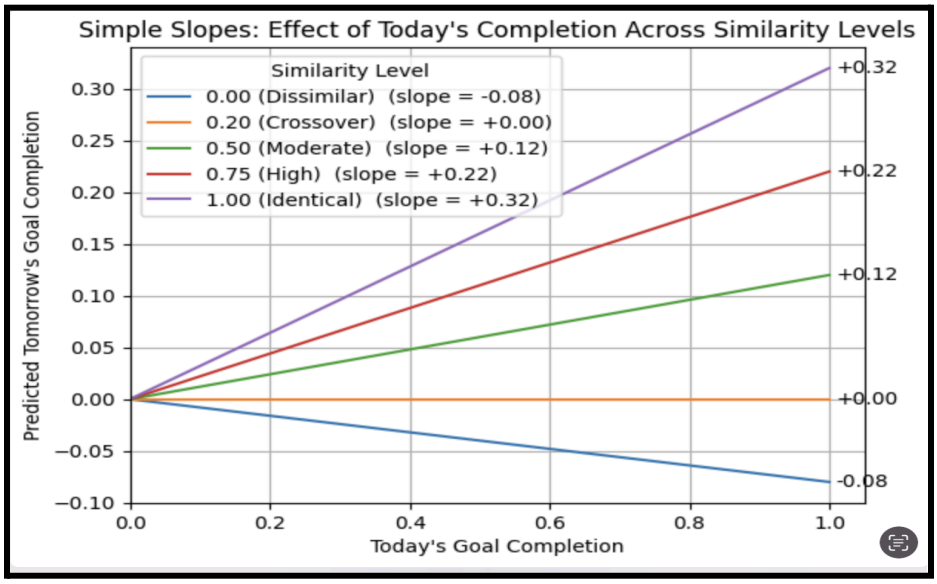


Figure 1: Effect of today's completion on tomorrow's completion at different levels of similarity:

Table 2: Predicting Goal Similarity from Today's Effort and Completion

Model: **similarity_cont $_{t \times [t+1]}$ ~ effort $_t$ x completion $_t$**

Predictor	β	SE	z	p	95% CI
Intercept	0.507***	0.020	25.29	<.001	[0.467, 0.546]
Today's Effort	0.020***	0.003	6.93	<.001	[0.015, 0.026]
Today's Completion	-0.001***	0.000	-3.66	<.001	[-0.001, -0.000]
Effort × Completion	-0.000077*	0.000039	-1.98	.047	[-0.000153, -0.000001]

LOOKING AT MODEL 1 & MODEL 2 (TABLE 1)

MAIN EFFECT OF GOAL SIMILARITY

When today's completion = 0 (didn't work on it), higher similarity predicts lower completion tomorrow. A plausible explanation: Difficult/effortful goals that weren't completed today get carried over as similar goals tomorrow, and they remain difficult. The negative main effect might reflect that ongoing challenging goals have similar content across days but low completion levels.

I want to look at the main effect more deeply. Plausible explanation: Difficult/effortful goals that weren't completed today get carried over as similar goals tomorrow, and they remain difficult.

To test this **(table 2)**, I checked if high effort and lower completion today → high similarity with tomorrow. From the results (table 2), we can see that higher effort goals today predict higher similarity with tomorrow. The negative interaction shows that the high effort → high similarity effect weakens as completion today increases. This makes sense: as we finish goals today, we're less likely to set similar goals tomorrow. But when completion is lower today, particularly for difficult goals, they get carried over to the next day.

INTERACTION EFFECT (Positive)

- If similarity between today and tomorrow's goals are high, then higher completion today leads to higher completion tomorrow. Or, if completion today is high and participants set a similar goal tomorrow, completion tomorrow is higher.
    - **ASIDE**: From **table 2**, we already see that participants are less likely to set a similar goal tomorrow if completion today is high. (We can extend this analysis by splitting the dataset into subgroups like recurring or habitual goals vs one off goals.) But this interaction term from model 2 (table 1) seems to indicate a momentum effect for high completion goal today and high similarity with tomorrow. This makes me feel there may beat least two subsets of the goal. One that has multi day spanning goals and another that sets domain similar goals across days when completion today is high.

- If completion today is high, participants may have a psychological carry-over: confidence, skill, or progress transfers. So a similar goal tomorrow may seem easier. This is genuine momentum. However, it might also be the case that the same goal is spanning multiple days (high similarity) and obviously as participants work towards the goal, with time, goal completion increases. So it might be good to isolate goals which are not spanning multiple days and recheck this interaction.

Therefore, I will isolate goals that span multiple days with goals that do not. **Continuation goal**s are ones where the completion for today is below 90th percentile for this participant and the goal is labelled as same as tomorrow by the LLM and goal similarity is greater than 0.80. Non Continuation goals are the remaining. I will repeat the analysis for both these datasets.

CONTINUATION VS NON CONTINUATION GOALS

- `Total observations: 14660`
- `Same as tomorrow: 3502 (23.9%)`
- `Low completion (<90th percentile): 10117 (69.0%)`
- `Continuation goals: 2793 (19.1%)`
- `Non-continuation goals: 11867 (80.9%)`

Continuation goals

	P	Date	Today's Goal	Tomorrow's Goal	Comp%	Sim
0	P1	02-01	Make bday present	Make bday	9	0.84
1	P1	02-05	Read bio	Read bio	73	1.00
2	P1	02-06	Read psych	Read psych	74	1.00
3	P1	02-11	Read psych	Read psych	10	1.00
4	P1	02-14	Do psych revision	Revise psych	78	0.87
5	P2	09-30	drink enough water	drink water	54	0.82
6	P2	10-13	finish bio readings	finish bio notes	97	0.80
7	P2	10-22	to watch two anthropology lectures	to watch all anthro lectures	76	0.82
8	P2	10-29	finish bio FLR	finish bio FLR	56	1.00
9	P2	10-30	finish bio FLR	finish bio FLR	83	1.00
10	P3	02-02	apply for volunteering	apply for volunteering	0	1.00
11	P3	02-05	bio lectures	bio lectures	70	1.00
12	P3	02-09	chem lectures	chem quiz	0	0.82
13	P3	02-10	chem quiz	chem quiz	0	1.00
14	P3	02-20	some exercising	some exercising	12	1.00

Non Continuation Goals

	P	Date	Today's Goal	Tomorrow's Goal	Matched Today Comp%	Sim
0	P1	09-29	Study for lab test	Finish co-op stuff	94	0.23
1	P1	09-29	Study for PSYB55	Hand in media assignment and do PSYB07 questions	51	0.59
2	P1	09-30	Hand in media assignment and do PSYB07 questions	Start cranial assignment comments	99	0.39
3	P1	09-30	Hand in media assignment and do PSYB07 questions	Review lecture 1-4 for PSYB55	99	0.54
4	P1	10-01	Review lecture 1-4 for PSYB55	Student PSYB55	0	0.63
5	P2	09-29	Eat healthy	Be healthy	59	0.82
6	P2	09-29	Eat healthy	Exercise	59	0.38
7	P2	09-30	Be healthy	Be healthy	100	1.00
8	P2	09-30	Exercise	Workout	99	0.66
9	P2	10-01	Workout	Workout	100	1.00
10	P3	01-31	finish half of my psych peer scholar assignment	wake up on time for my psych lecture	49	0.49
11	P3	01-31	participate in my biomed tutorial	be prepared for my bio lab	100	0.53
12	P3	02-01	be prepared for my bio lab	finish my peer scholar assignment	90	0.36
13	P3	02-01	be prepared for my bio lab	write pre-reading notes for astro	90	0.36
14	P3	02-02	finish my peer scholar assignment	complete post lab quiz for chem	65	0.33

TABLE 3

- Model 4 (**Only Non Cont Goals**): **completion $_{t+1}$ ~ similarity_binary $_{t \times [t+1]}$ x completion $_{t}$**
- Model 5 (**Only Non Cont Goals**): **completion $_{t+1}$ ~ similarity_cont $_{t \times [t+1]}$ x completion $_{t}$**

Predictor	Model 1: Binary (Same as Tomorrow)	Model 2: Continuous Similarity
Fixed Effects		
Intercept	**58.35*** (1.36)**	**62.96*** (2.12)**
Today's Completion	**0.05*** (0.01)**	**-0.06* (0.02)**
Same as Tomorrow (Binary)	**-5.15** (1.70)**	**—**
Similarity (Continuous)	**—**	**-13.71*** (3.99)**
Interaction		
Today's Completion × Similarity	**0.14*** (0.03)**	**0.34*** (0.06)**
Model Fit		
N observations	**11,046**	**11,046**
N groups	**112**	**112**

Looking at the interaction effects of both models, if goals are highly similar, higher completion today predicts higher completion tomorrow.

Since these are non continuous goals, this means the high similarity goals within this subset may either be habitual goals or goals under the same domain. I repeated the analysis restricting to non-continuation goals with similarity below 0.70, 0.60, and 0.50. The momentum effect remained significant at all thresholds

TABLE 4

- **$\text{completion}_{t+1} \sim \text{similarity\_cont}_{t \times [t+1]} \times \text{completion}_{t} + \text{pos\_affect}_{t+1}$**
- **Similar coefficients for non continuation goals**

Model	Momentum Interaction (β)	p-value	Affect Effect (β)	Change
Without Affect	0.400***	<.001	—	—
With Affect	0.392***	<.001	6.813***	**-2.0%**

```
Momentum Effect (Completion × Similarity Interaction):
  Without affect: β = 0.4002, p = 0.0000
  With affect:    β = 0.3921, p = 0.0000

  Change: 2.0% reduction
```

Affect does NOT mediate momentum:

- Adding tomorrow's positive affect to the model reduces the momentum effect by only **2.0%**
- The momentum interaction remains virtually unchanged (0.400 → 0.392)
- However, affect independently predicts completion (β = 6.813, p < .001)

COMMENTS
Misc Comments

The key predictor is `prev_goal_completion`, defined as the average completion rate of the two goals set on day *t*, while the outcome variable is `goal_sim`, the maximum similarity between goals on days *t* and *t + 1*.

I tested if high effort today → lower completion today. This relationship was not significant.

TBD: REPEAT WITH

References

REFERENCE

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, 56(3), 218-226.

Lally, Phillippa & Jaarsveld, Cornelia & Potts, Henry & Wardle, Jane. (2010). How are habits formed: Modeling habit formation in the real world. European Journal of Social Psychology. 40. 10.1002/ejsp.674.

Louro, Maria & Pieters, Rik & Zeelenberg, Marcel. (2007). Dynamics of Multiple-Goal Pursuit. Journal of Personality and Social Psychology. 93. 174-193. 10.1037/0022-3514.93.2.174.

Fishbach, A., Zhang, Y., & Koo, M. (2009). The dynamics of self-regulation. European Review of Social Psychology, 20(1), 315–344. https://doi.org/10.1080/10463280903275375