Supplemental Material

Table of Contents	Page Number
Supplemental Study with Experimental Manipulation of Scale Training	1
Table S1: Perceived Naturalness of Additive versus Subtractive Beverages by Scale Training in Supplemental Study	3
Study 1 Supplemental Results	4
Table S2: Perceived Naturalness of Additive versus Subtractive Beverages in Study 1	4
Study 3 Supplemental Results	5
Table S3: Perceived Naturalness of Additive versus Subtractive Framed Items in Study 3	5
Study 4A Supplemental Results	6
Table S4: Perceived Naturalness of Items Described with "Fortified" versus "Additive" in Study 4A	6
Table S5: Perceived Naturalness of Items Described with "Supplemented" versus "Additive" in Study 4A	7
Table S6: Perceived Naturalness of Items Described with "Supplemented" versus "Fortified" in Study 4A	7
Study 4B Supplemental Results	8
Table S7: Ratings of Examples of Fortifiers versus Additives on Valence, Healthfulness, and Naturalness in Study 4B.	8
Table S8: Ratings of Examples of Supplements versus Additives on Valence, Healthfulness, and Naturalness in Study 4B.	9
Table S9: Ratings of Examples of Supplements versus Fortifiers on Valence, Healthfulness, and Naturalness in Study 4B.	9
Study 5 Supplemental Results	10
Table S10: Perceived Naturalness of Pairings where Dosage is Tripled in Study 5.	10
Table S11: Perceived Naturalness of Pairings where Amount of Processing is Tripled in Study 5.	11
Reference	12

Supplemental Study with Experimental Manipulation of Scale Training

We ran a study to assess whether the presence of scale training/instructions (see Study 1 Methods in main text) altered patterns of naturalness ratings and additivity dominance. We did not expect scale training to affect additivity dominance, because training and examples do not highlight or distinguish between additives and subtractives.

Method

We used the stimuli from Study 1, which examined perceived naturalness of additive and subtractive beverage pairs (e.g., milk with fat added versus milk with fat removed). We made two changes to Study 1's methods. First, we manipulated the presence of scale training to understand whether training on how to use the naturalness scale altered naturalness judgments. We did not expect scale training to affect our results. Second, we randomized which block was presented first, orange juice beverages or milk beverages (whereas in the original study participants always rated milk beverages and then orange juice beverages).

Three hundred and five participants completed a web-based survey on Amazon's Mechanical Turk in September 2017 (48.2% female, M_{age} = 35.8, SD = 11.1). Note that because half of participants did not complete any scale training, we do not make any exclusions based on scale training.

Results

We conducted a 2 (Process: Added, Subtracted) X 5 (Beverage Pairs: Milk with Fat, Milk with Sugar, Milk with Calcium, Orange Juice with Pulp, Orange Juice with Vitamin C) X 2 (Scale Training: Yes, No) mixed ANOVA on the rated naturalness of the five additive-subtractive beverage pairs (ten beverages total). The first two factors were between-subjects and the last factor was within-subjects. Consistent with additivity dominance, there was an effect of

process such that additive beverages were generally rated less natural than subtractive beverages ($F(1, 303) = 37.87$, $p < 0.001$, $\eta_p^2 = 0.11$). Additionally, product pairs varied in naturalness ($F(4, 1212) = 49.92$, $p < 0.001$, $\eta_p^2 = 0.14$). Additivity dominance was larger for some beverage pairs (Process by Beverage Pair interaction: $F(4, 1212) = 10.85$, $p < 0.001$, $\eta_p^2 = 0.04$). The main effect of scale training was not reliable ($F(1, 303) = 1.89$, $p = .17$) nor were interactions with scale training (Scale Training X Process interaction: $F(1, 303) = .57$, $p = .452$; Scale Training X Beverage Pair interaction: $F(1, 303) = 2.00$, $p = .092$, $\eta_p^2 = 0.01$; Scale Training X Process X Beverage Pair interaction: $F(1, 1212) = 2.28$, $p = .058$, $\eta_p^2 = 0.01$). Naturalness of additive versus subtractive beverages are displayed in Table S1, broken down by whether or not participants received the scale training. When participants received scale training, additive beverages were rated as significantly less natural than subtractive beverages for four out of five pairings (all except milk and calcium). When participants did not receive scale training, additive beverages were rated as significantly less natural than subtractive beverages for three out of five pairings. Surprisingly, in this condition one pair showed significant results in the opposite direction: milk with calcium added was rated as more natural than milk with calcium removed when there was no scale training.

Discussion

In a supplemental study, we directly replicate results from Study 1. We also rule out the possibility that additivity dominance only occurs when participants receive our scale training instructions.

Table S1

Perceived Naturalness of Additive versus Subtractive Beverages by Scale Training in

Supplemental Study

	Scale Training	Mean Naturalness of Additive Product (S.D.)	Mean Naturalness of Subtractive Product (S.D.)	T-Value	P-Value	Cohen's d_{RM}
Milk with Fat	Yes	49.68 (26.28)	55.51 (26.14)	3.44	<.001	0.28
Milk with Sugar	Yes	44.01 (25.32)	51.35 (25.37)	5.00	<.001	0.4
Milk with Calcium	Yes	51.15 (24.81)	52.28 (24.70)	.83	.409	0.07
Orange Juice with Pulp	Yes	56.67 (25.73)	61.56 (26.23)	2.46	.015	0.2
Orange Juice with Vitamin C	Yes	52.17 (25.85)	57.70 (26.88)	3.29	.001	0.26
Milk with Fat	No	53.67 (27.00)	56.76 (27.72)	1.63	.105	0.13
Milk with Sugar	No	47.54 (26.47)	55.62 (25.77)	4.69	<.001	0.38
Milk with Calcium	No	54.74 (26.24)	49.92 (25.65)	-2.85	.005	-0.23
Orange Juice with Pulp	No	61.09 (27.98)	68.84 (25.21)	4.01	<.001	0.33
Orange Juice with Vitamin C	No	55.99 (25.20)	61.23 (25.44)	2.75	.007	0.23

Note. Mean naturalness ratings on 0 to 100 scale are displayed with standard deviations in

parentheses in columns 2 and 3. Because the experimental design is within-subjects, t-tests are

paired sample t-tests and effect sizes (Cohen's d_{RM}) are the mean difference score divided by the

standard deviation of the difference scores (see Morris & DeShon, 2002).

Study 1 Supplemental Results

Table S2 displays the descriptive statistics and t-tests for each additive-subtractive pairing in Study 1. Consistent with the additivity dominance for beverages hypothesis, additive products are rated as less natural than subtractive products.

Table S2

Perceived Naturalness of Additive versus Subtractive Beverages in Study 1

	Mean Naturalness of Additive Product (S.D.)	Mean Naturalness of Subtractive Product (S.D.)	T-Value	P-Value	Cohen's d_{RM}
Milk with Fat	54.87 (24.41)	60.77 (22.01)	3.74	<.001	0.30
Milk with Sugar	54.36 (24.10)	58.26 (23.07)	2.83	0.005	0.23
Milk with Calcium	56.62 (23.93)	58.96 (22.11)	1.65	0.101	0.13
Orange Juice with Pulp	65.28 (22.32)	68.41 (22.43)	2.00	0.047	0.16
Orange Juice with Vitamin C	62.10 (21.91)	66.55 (21.11)	2.79	0.006	0.23

Note. Mean naturalness ratings on 0 to 100 scale are displayed with standard deviations in parentheses in columns 2 and 3. Because the experimental design is within-subjects, t-tests are paired sample t-tests and effect sizes (Cohen's d_{RM}) are the mean difference score divided by the standard deviation of the difference scores (see Morris & DeShon, 2002).

4

Study 3 Supplemental Results

Table S3 displays the descriptive statistics and t-tests for each additive-subtractive frame pairing in Study 3. Consistent with the additivity dominance framing hypothesis, items are rated as less natural when they are framed as additives versus subtractives.

Table S3

Perceived Naturalness of Additive versus Subtractive Framed Items in Study 3

	Mean Naturalness of Additive Framing (S.D.)	Mean Naturalness of Subtractive Framing (S.D.)	T-Value	P-Value	Cohen's d_{RM}
Peanut Oil	52.84 (28.69)	60.46 (29.96)	4.39	<.001	0.32
Yogurt Fat	49.89 (28.42)	57.90 (29.76)	4.62	<.001	0.33
Pulp	61.17 (27.49)	71.92 (26.18)	6.32	<.001	0.46
Calcium	48.65 (29.90)	57.23 (30.96)	4.70	<.001	0.34

Note. Mean naturalness ratings on 0 to 100 scale are displayed with standard deviations in parentheses in columns 2 and 3. Because the experimental design is within-subjects, t-tests are paired sample t-tests and effect sizes (Cohen's d_{RM}) are the mean difference score divided by the standard deviation of the difference scores (see Morris & DeShon, 2002).

Study 4A Supplemental Results

Tables S4-S6 display the descriptive statistics and t-tests for perceived naturalness of items depending on the synonym ("additive", "fortified" or "supplemented") used to describe the item. The connotation account hypothesis predicts items described with "additive" will be rated as less natural, because additive is more negatively valenced than fortified or supplemented. Contrary to this hypothesis, additive items were directionally rated as more natural than fortified items (Table S4) and supplemented items (Table S5). Supplemented and fortified items did not differ in naturalness (see Table S6).

Table S4

Perceived Naturalness of Items Described with "Fortified" versus "Additive" in Study 4A

Item	Mean Naturalness of Fortified Description (S.D.)	Mean Naturalness of Additive Description (S.D.)	T-Value	P-Value	Cohen's D
Orange Juice with Pulp Added	52.15 (26.86)	59.66 (23.94)	1.72	.087	-.30
Peanut Butter with Fat Added	39.05 (25.74)	47.80 (25.29)	2.00	.048	-.34
Milk with Calcium Added	48.46 (24.72)	54.31 (25.73)	1.35	.180	-.23

Note. Mean naturalness ratings on 0 to 100 scale are displayed with standard deviations in parentheses in columns 2 and 3, and between-subjects t-tests comparing the two descriptions are displayed in columns 3-5.

Table S5

Perceived Naturalness of Items Described with "Supplemented" versus "Additive" in Study 4A

Item	Mean Naturalness of Supplemented Description (S.D.)	Mean Naturalness of Additive Description (S.D.)	T-Value	P-Value	Cohen's D
Orange Juice with Pulp Added	55.04 (24.83)	59.66 (23.94)	1.07	.287	-.19
Peanut Butter with Fat Added	45.19 (23.05)	47.80 (25.29)	.60	.547	-.11
Milk with Calcium Added	53.72 (21.85)	54.31 (25.73)	.14	.891	-.03

Note. Mean naturalness ratings on 0 to 100 scale are displayed with standard deviations in parentheses in columns 2 and 3, and between-subjects t-tests comparing the two descriptions are displayed in columns 3-5.

Table S6

Perceived Naturalness of Items Described with "Supplemented" versus "Fortified" in Study 4A

Item	Mean Naturalness of Supplemented Description (S.D.)	Mean Naturalness of Fortified Description (S.D.)	T-Value	P-Value	Cohen's D
Orange Juice with Pulp Added	55.04 (24.83)	52.15 (26.86)	.61	.542	.11
Peanut Butter with Fat Added	45.19 (23.05)	39.05 (25.74)	1.38	.170	.25
Milk with Calcium Added	53.72 (21.85)	48.46 (24.72)	1.24	.218	.23

Note. Mean naturalness ratings on 0 to 100 scale are displayed with standard deviations in parentheses in columns 2 and 3, and between-subjects t-tests comparing the two descriptions are displayed in columns 3-5.

Study 4B Supplemental Results

Tables S7-S9 display the descriptive statistics and t-tests for valence, naturalness, and healthfulness ratings of participant-generated examples of additives, fortifiers, and supplements. Additive examples are rated as less positively valenced, less natural, and less healthy than fortifier examples (Table S7) and supplement examples (Table S8). Valence, naturalness, and healthfulness of fortifier and supplement examples do not reliably differ (Table S9). This pattern of results is consistent with an account where different synonyms elicit different psychologically accessible examples of the entity to be added. When we control for these different denotations by specifying which entity has been added in Study 4A, connotations of the synonyms do not affect naturalness ratings.

Table S7

Ratings of Examples of Fortifiers versus Additives on Valence, Healthfulness, and Naturalness in Study 4B.

Dimension	Mean Rating of Fortifier Example (S.D.)	Mean Rating of Additive Example (S.D.)	T-Value	P-Value	Cohen's d_{RM}
Valence	2.01 (1.44)	-0.68 (1.97)	16.35	<.001	1.15
Healthfulness	1.94 (1.49)	-1.45 (1.48)	23.89	<.001	1.68
Naturalness	1.82 (1.51)	-0.70 (1.90)	15.52	<.001	1.09

Note. In columns 2 and 3, mean ratings on -3 to 3 scale for valence, healthfulness, and naturalness of different examples are displayed with standard deviations in parentheses. Because the experimental design is within-subjects, t-tests are paired sample t-tests and effect sizes (Cohen's d_{RM}) are the mean difference score divided by the standard deviation of the difference scores (see Morris & DeShon, 2002).

Table S8

Ratings of Examples of Supplements versus Additives on Valence, Healthfulness, and

Naturalness in Study 4B.

Dimension	Mean Rating of Supplement Example (S.D.)	Mean Rating of Additive Example (S.D.)	T-Value	P-Value	Cohen's d_{RM}
Valence	1.97 (1.45)	-0.68 (1.97)	15.82	<.001	1.11
Healthfulness	1.98 (1.46)	-1.45 (1.48)	24.26	<.001	1.71
Naturalness	1.59 (1.68)	-0.70 (1.90)	13.26	<.001	.93

Note. In columns 2 and 3, mean ratings on -3 to 3 scale for valence, healthfulness, and

naturalness of different examples are displayed with standard deviations in parentheses. Because

the experimental design is within-subjects, t-tests are paired sample t-tests and effect sizes

(Cohen's d_{RM}) are the mean difference score divided by the standard deviation of the difference

scores (see Morris & DeShon, 2002).

Table S9

Ratings of Examples of Supplements versus Fortifiers on Valence, Healthfulness, and

Naturalness in Study 4B.

Dimension	Mean Rating of Supplement Example (S.D.)	Mean Rating of Fortifier Example (S.D.)	T-Value	P-Value	Cohen's d_{RM}
Valence	1.97 (1.45)	2.01 (1.44)	.36	.722	.03
Healthfulness	1.98 (1.46)	1.94 (1.49)	-.31	.758	-.02
Naturalness	1.59 (1.68)	1.82 (1.51)	1.76	.081	.12

Note. In columns 2 and 3, mean ratings on -3 to 3 scale for valence, healthfulness, and

naturalness of different examples are displayed with standard deviations in parentheses. Because

the experimental design is within-subjects, t-tests are paired sample t-tests and effect sizes

(Cohen's d_{RM}) are the mean difference score divided by the standard deviation of the difference

scores (see Morris & DeShon, 2002).

Study 5 Supplemental Results

Table S10 displays descriptive statistics and t-tests for naturalness ratings when the dosage of an additive is tripled (holding amount of processing constant). Table S11 displays the descriptive statistics and t-tests for naturalness ratings when the amount of processing is tripled (holding the dosage of the additive constant). Consistent with the extra processing account, tripling the processing significantly reduces naturalness. However, consistent with the contagion hypothesis, the effects of tripling dose and tripling processing are both small.

Table S10

Perceived Naturalness of Pairings where Dosage is Tripled in Study 5.

Dimension	Mean Rating of 5% via 1 Process (S.D.)	Mean Rating of 15% via 1 Process (S.D.)	T-Value	P-Value	Cohen's d_{RM}
Milk with Fat	60.03 (26.26)	56.68 (26.98)	3.38	<.001	.24
Milk with Calcium	59.31 (25.46)	57.52 (25.94)	1.80	.073	.13
Pulp with Orange Juice	66.58 (25.86)	60.64 (27.61)	4.61	<.001	.33

Note. Mean naturalness ratings on 0 to 100 scale are displayed with standard deviations in parentheses in columns 2 and 3. Because the experimental design is within-subjects, t-tests are paired sample t-tests and effect sizes (Cohen's d_{RM}) are the mean difference score divided by the standard deviation of the difference scores (see Morris & DeShon, 2002).

Table S11

Perceived Naturalness of Pairings where Amount of Processing is Tripled in Study 5.

Dimension	Mean Rating of 15% via 1 Process (S.D.)	Mean Rating of 15% via 3 Processes (S.D.)	T-Value	P-Value	Cohen's d_{RM}
Milk with Fat	56.68 (26.98)	51.02 (28.44)	4.46	<.001	.32
Milk with Calcium	57.52 (25.94)	50.61 (27.47)	5.87	<.001	.42
Pulp with Orange Juice	60.64 (27.61)	59.17 (28.55)	1.23	.220	.09

Note. Mean naturalness ratings on 0 to 100 scale are displayed with standard deviations in parentheses in columns 2 and 3. Because the experimental design is within-subjects, t-tests are paired sample t-tests and effect sizes (Cohen's d_{RM}) are the mean difference score divided by the standard deviation of the difference scores (see Morris & DeShon, 2002)

Reference

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*(1), 105-125.