## Supplemental Online Material

## Study 1

*Notes. In the following section, we provide details on the development of the word lists used in the semantic coherence task (see ‘Study 1- Semantic coherence task’ in the original manuscript).*

## Method

*Semantic Coherence Task.* To examine the influence of thematic compared to taxonomic relations on people’s intuitive judgments of semantic coherence, we developed a new set of word triads that contained either a thematic or a taxonomic relation. For this purpose, we primarily relied on the word list of Fröhlich and Högl (2012). The word list was originally developed to assess people’s preference for thematic over taxonomic relations (Lin & Murphy, 2001). In the original preference task, participants were instructed to choose one out of two words that was more related to a target word (*e.g., DOG)*. While one word was primarily thematically related to the target word (*e.g., BONE*), the other word was primarily taxonomically related (*e.g., CAT*). Previous studies examined the validity of the thematic preference task by controlled for confounding factors such as the associative strength between the target word and their thematic and taxonomic relation (e.g. Fröhlich & Högl, 2012). We adapted the task out of two different reasons. First, several words in the lists were derived from an U.S.-cultural background. For instance, the original list contained the triad HAWAI, MISSOURI and BEACH. American students might have identified Hawaii and Missouri as U.S. territories with ease. Germans probably would not. Thus, we adapted the words to fit to the cultural context. Second, previous findings consistently found that people preferred thematic relations (e.g. Fröhlich & Högl, 2012). Although unlikely, these differences might originate from differences in associative strength or semantic relatedness. To rule out such differences, we chose new thematically related words for some trials if these words appeared to be highly semantically related, or adapted the taxonomic words if these were weakly related to the targets. As documented below in Study 2, this procedure was successful, as we reduced people’s preference for thematic relations to chance level.

## Study 2

*Notes. In the supplemental material of Study 2, we include a detailed description of the thematic preference task that was included in Study 2 for exploratory research, but excluded from the original manuscript (see Study 2- Procedure). Moreover, in the original manuscript (Study 1-Semantic Coherence Task), we highlighted that the thematic and taxonomic words of this task are equally related to the target in terms of semantic relatedness. In the Results section of this supplemental material, we provide results supporting this notion.*

*In the original manuscript (Study 2), we noted that we assessed mood. Results on mood are presented in the Results section of this supplemental material. Finally, an anonymous reviewer asked us to report results on response times of intuitive coherence judgments in this supplemental material. Results are documented in the Results section as well.*

In Study 2, we chose a novel methodological approach to investigate the influence of thematic compared to taxonomic processing in intuitive judgments of semantic coherence. For exploratory purpose, we further included a thematic preference task to examine the impact of the manipulation on people’s semantic processing. We chose the thematic preference task as a non-intuitive judgment task that examines semantic processing (Fröhlich & Högl, 2012). As this research question was unrelated to the judgments of semantic coherence, we dropped it from the original manuscript. Yet, for exploratory purpose, we decided to include the results in the supplemental material.

*Thematic preference task.* We assessed the preference for thematic over taxonomic using a validated task (Fröhlich & Högl, 2012). In this test, participants are instructed to choose which from two words is more related to a target word (*e.g., DOG)*. One word was thematically related (*e.g., BONE*), the other word is taxonomically related (*e.g., CAT*). We adapted the original word list to balance people’s preference for thematic or taxonomic relations (see Study 1). Participants were introduced to 39 target words. Note that half of the target words and their related words were used in the anagram task as well. The related words were presented below the target word. Participants were instructed to spontaneously choose either of the two related words that appeared to be more related to the target word pressing either ‘x’ for the left word or ‘m’ for the right word. We pseudo-randomly balanced words’ position. Instructions stressed that words’ relatedness was highly subjective and participants were free to choose the word they preferred within ten seconds.

## Results

*Thematic preference task.* First, we controlled for participants’ bias for thematic over taxonomic relations. As outlined above, we adapted the preference task to balance people’s preference for thematic over taxonomic relations. Correspondingly, a one-sample t-test confirmed that participants did not prefer thematic relation more than chance level, *M* = .50, *SD* = .18, *p*= .855.

Second, we examined the influence of the priming task on people’s preference for thematic relations. For this purpose, we conducted an ANOVA with thematic priming as between-participant factor and material (familiar versus unfamiliar trials) as within-participant factor. Neither thematic priming nor material yielded a significant influence of people’s preferences, *ps* ≥ .189. Still, a significant interaction emerged, *F(1, 85)*= 16.20, *p*< .001, η² = .16. Bonferroni-correct post-hoc t-tests indicated that participants chose those words as more related to the target word that they encountered previously in the priming task, *t(85)*= 2.62, *p* < .01. However, we found no priming effect for trials that were not used in the priming task, *p*= .669.

We repeated the analysis with people’s reaction time as independent variable. Both triads and material yielded a significant influence on response times, *Fs(1, 85)*≥ 6.53, *p*< .05, η² ≥ .07. Participants chose faster for familiar triads, *Mfamiliar*= 1864.41, *SDfamiliar* = 56.27, compared to unfamiliar triads, *Munfamiliar*= 1983.34, *SDunfamiliar* = 60.11. Moreover, participants chose faster in the thematic priming condition, *Mthematic*= 1777.40, *SDthematic* = 79.65, compared to the taxonomic priming condition, *Mtaxonomic*= 2070.25, *SDtaxonomic* = 82.45.

Taken together, findings suggest that the priming task affected participants’ judgment process in the thematic preference task, but did not influence people’s final preference judgment, at least for those words that were not presented previously. Findings indicate that a thematic priming condition eased participants’ judgment process. In the anagram task, we found that participants reported more difficulties to grasp the semantic nature of the thematically related anagram word. Thus, participants might had difficulties deliberately grasping the relatedness of the thematically related words in the preference task. Encountering thematic relations previously may have helped participants to grasp thematic relations faster in the preference task. Yet, an ease in thematic processing did not lead to a shift in people’s deliberate preference judgments.

*Mood.* To explore the influence of the priming procedure on mood, we conducted a MANOVA with positive and negative mood as dependent variables. We found no significant influence of priming on mood, *ps* ≥ .493.

Moreover, Pearson-correlations found no significant relationship between positive and negative mood and intuitive coherence judgments’ accuracy and bias, *ps* ≥ .159.

*Response times of intuitive coherence judgments.* We conducted a repeated measure ANOVA for response times in intuitive coherence judgments. The 2 (coherent vs incoherent triad) x 2 (coherent vs. incoherent judgment) x 2 (thematic vs. taxonomic priming) ANOVA indicated no significant influence of priming, *p* = 462. We only found a significant difference between judgment speed for coherent compared to incoherent triads, *F(1, 85)*= 6.38, *p*< .05.Incoherent triads were judged faster compared to coherent triads, *Mcoh*= 1622.30, *SDcoh*=40.71, *Mincoh*= 1599.16, *SDcoh*=40.44.

## Study 3

*Notes. In the original manuscript, we presented a summary of the original datasets used in Study 3 (see Study 3- Original Datasets). In this supplemental material, we present the datasets in detail.*

***Original Datasets***

*Dataset A.* The first dataset consisted of two published experiments by Bolte et al. (2003). The specific version of the semantic coherence task in this study consisted of 36 triads with, and 36 triads without common denominator derived from the original set (Bowers et al., 1990). Participants had three seconds to read and evaluate each triad. Afterwards, the entire set of triads was presented again and participants were instructed to search for a potential denominator. For the analysis, triads were excluded when the correct denominator was detected. In total, we obtained coherence judgments for the triads assessed in three different experimental conditions. In the first two conditions, all triads were administered. In the third condition, only 24 triads were used. To ensure that coherence judgments of the word triads were stable, we calculated Pearson correlations. Results indicated that intuitive coherence judgments were rather stable, *r(72)*= .57, *p* < .001 to *r(24) =*.37, *p* = .08. For the sake of simplicity, we therefore calculated average coherence judgments for each triad.

*Dataset B.* The second dataset contained results from two published studies by Kuhl and Kazén (2008) and Kazén, Kuhl, and Quirin (2015). In both studies, the semantic coherence task was administered in the same way as in Dataset A. The stimulus set consisted of 24 triads with a common denominator, and 24 triads without a common denominator. Most of them were derived from the triad set used by Bolte and colleagues (Bolte et al., 2003; Bolte & Goschke, 2005); still, 25 triads differed from the original set. We calculated Pearson correlation to investigate the stability of the coherence judgments of word triads. Coherence judgments between the two studies were highly stable, *r(48)* = .92, *p* < .001. Thus, we computed average coherence judgments for each triad. Note that explicit solutions were not excluded in this dataset.

*Dataset C.* The third dataset originated from an unpublished study by Schönbrodt and Hauser (2015). In this study, the semantic coherence task consisted of 50 triads with a common denominator, and 50 triads without a common denominator Schönbrodt and Hauser revised the original set of word triads that was published by Bowers and colleagues (1990). Although several triads were similar to those of Bolte and colleagues (2003), Schönbrodt and Hauser (2015) made slight adaptations for most triads. The first purpose of their work was to translate the triads into modern German. Second, they aimed to develop a more difficult set of triads in which the triads’ hidden denominator could not be found explicitly. To create triads without a common denominator, triad words were ordered by pseudo-random reassignment. In the latter version of the semantic coherence task, participants first read the triads for 1.5 seconds. Next, participants had 2 seconds to judge the triads in terms of coherence. Whenever participants judged the triad as coherent, they were asked to name a potential solution word (compare Bolte & Goschke, 2005; Topolinski & Strack, 2008; 2009). For this dataset, we calculated two different scores for each triad. For the first score (Cstrict), we applied the conventional approach in intuition research (e.g. Baumann & Kuhl, 2002; Bolte et al., 2003) and excluded all triads for which participants had found the correct solution word from further analysis. For the second score (Clax), we included explicitly solved triads similar to Dataset B. Thus, the dataset includes both intuitive and non-intuitive judgments of semantic coherence.

## Study 3d

*Notes. In the manuscript, we excluded details on the instructions and results of manipulation checks of Study 3d due to word limitation (see Procedure and materials section of Study 3). We included detailed information in the subsequent sections. Moreover, in the Discussion section of this supplemental online material, we provided a detailed discussion on potential reasons why the present findings on thematic relatedness may deviate from previous studies. We shortly discussed this issue in the Discussion of Study 3 in the original manuscript.*

*Instructions.* In the thematic relations condition, we instructed participants to estimate the thematic relatedness between the word pairs. Thematic relations were defined as a type of relations that describe the extent to which things go together in the same scenario or event. In the taxonomic relations condition, taxonomic relations were defined as the extent to which things share the same features and thus are part of the same category. We presented and explained five examples to illustrate the different types of thematic (taxonomic) relations including noun-noun, noun-verb, noun-adjective and verb-verb combinations. Next, we reminded participants not to rely on other sorts of associations such as taxonomic (or thematic) relations. To examine the quality of the instructions, we asked participants responded whether they have understood what thematic (or taxonomic) relations were (0 = *‘yes’*, 1 = *‘no’*). Next, we presented participants seven word pairs and asked them to choose, which of those word pairs was thematically (or taxonomically) related. Then, we explained each example. Examples ranged from word pairs that were solely thematically *or* taxonomically related, as well as pairs that were both thematically and taxonomically related to different extents. Finally, at the end of the training phase, participants rated on a four-point Likert scale how difficult it was to rate the respective type of relations in the manipulation check examples (0= *‘not difficult at all’* to 4= *‘very difficult’*)

## Results

First, we examined how difficult it was to rate the thematic and taxonomic pairs presented in the instructions. A t-test revealed that participants had more difficulties rating taxonomic compared to thematic relations in the pairs presented in the instructions, *Mthem*= 2.18, *SDthem* = .91, *Mtax*= 2.54, *SDtax* = .80, *T (99)* = 2.086, *p* < .05.

The difficulties in judging the triads might have influences participants’ final judgments of relatedness. However, Pearson-correlations revealed no significant relationship between rating difficulties and ratings on taxonomic or thematic relations, *ps*≥ .140. Moreover, when we repeated the ANOVA presented in the manuscript with repeated measure on presence of a common denominator and the between-participant factor relations as between factor and included rating difficulties as a control variable, we found no influence of ratings difficulty or an interaction effect on people’s ratings, *p* ≥.671. Both presence of a common denominator, relations and the interaction between the two factors remained significant, *FS (1, 98)* ≥ 39.204, *p* < .001.

**Discussion**

In Study 3, we found that triads with versus without a hidden denominator differed in terms of thematic relatedness.Results are seemingly at odds with previous published studies that found no differences between triads with versus without a common denominator in terms of semantic relatedness (Bolte & Goschke, 2005; Topolinksi & Strack, 2009). A closer inspection of these previous studies revealed, however, that the previous findings were somewhat ambiguous.

Specifically, both Bolte and Goschke’s (2005) and Topolinski and Strack’s (2009) observed ratings of semantic relatedness were in the direction predicted by the thematic integration model. Word pairs of coherent triads were rated as more semantically related than word pairs of incoherent triads. Specially, Topolinski and Strack (2009) reported an average relatedness of 2.38 (*SD* =.70) for coherent and 2.29 (*SD* =.88) for incoherent triads in a sample of twenty participants (*ts ≤*.05). Bolte and Goschke (2005) collected data from forty participants and found comparable relatedness ratings. However, due to the larger sample size, statistical differences were stronger (*t(39)* = -1.31*, p >*.19). Thus, the lack of statistical differences in semantic relatedness may partly originate from an underpowered design.

Moreover, it is important to note that neither Bolte and Goschke (2005) nor Topolinski and Strack (2009) were particularly interested in thematic relations. It is therefore conceivable that these researchers used measures and procedures that were not optimally suited to detect a difference in thematic relatedness between words from coherent versus incoherent triads.

To gain more insight into this matter, we emailed these researchers during October, of 2018. They could not provide with their actual materials, which is understandable given that the relevant studies were conducted over a decade ago (mail communication, 2018). However, the researchers did provide some useful background information that. Specifically, Bolte indicated that she asked participants to rate how much the words were associated. These instructions were sufficient to obtain a rough estimate of associative strength or semantic relatedness between the words that constitute the triads. However, as explained in the introduction, mere associations are theoretically distinct from thematic relations. Thus, Bolte’s instructions were not optimal for tapping into the thematic relatedness of the triad words. In a similar vein, Topolinski indicated that he might have first explained to his participants what semantic relations are. These instructions could have diminished the influence of thematic relations on people’s judgments of relatedness, given that thematic relations are but one type of semantic relations.

In view of these considerations, it appears that prior published findings on the semantic coherence task are not informative about potential differences in thematic relatedness, between words from coherent versus incoherent triads.

## Study 4

*Notes. In the following sections, we will outline the development of the triad lists. We dropped details on the procedure from the original manuscript due to word limit (see Study 4-Method). Moreover, we present an additional analysis on the influence of thematic relatedness and presence of a common denominator on people’s intuitive coherence judgments, as recommended by an anonymous reviewer.*

*Materials.* We created two versions of the semantic coherence task. In the original version of the task, triads with (versus without) a common denominator were more thematically related. To create an alternate version of the task, we used the thematic relation norms. We selected the triads in a way that allowed a maximum amount of word triads while maintaining a significantly negative relationship between presence of common denominator and thematic relations. To this end, we used the 18 most thematically related triads without hidden denominator and the 18 least thematically related triads that still contained a denominator. For the control list, we randomly chose 36 word triads with half of them containing a common denominator.

An independent t-test with presence of denominator as independent variable and thematic relations as dependent variable confirmed that triads without a common denominator were significantly more thematically related than triads with a common denominator *t( 28.31)* = 3.10, *p* < .01, *Mno denominator* = .83, *SDno denominator*= .26, *Mdenominator* = .61, *SDdenominator* = .16. For the control list, we randomly chose 18 triads without denominator and 18 triads with denominator from the data pool. An independent t-test confirmed that triads without a common denominator were significantly less thematically related than those triads that contain a common denominator, *t(26.63)*= -5.80, *p* < .001, *Mno denominator* = .48, *SDno denominator* = .16, *Mdenominator* = 1.01, *SDdenominator*= .36.

## Results

*Additional analysis.*Results of Study 4 indicate that the difference in intuitive coherence judgments of triads with compared to without common denominator was much larger in the standard list compared to the alternate list. Regarding this result, an anonymous reviewer wondered how these differences may be explained by thematic integration model. .

We suggested that the differences in intuitive coherence judgments may originate from larger differences in thematic relatedness between triads with versus without common denominator in the standard list compared to the alternate list. Consistently, we found that the presence of a common denominator and thematic relatedness were positively related across lists, *r (46)*= .34, *p*< .05. Next, we conducted an additional regression analysis predicting intuitive coherence judgment across the two lists, including presence of a common denominator and thematic relatedness as predictors. Consistent with Study 3, we found that thematic relatedness predicted intuitive coherence judgments, β = .55, *t(44)*= 4.11, *p* < .001. At the same time, presence of a common denominator did no longer predict coherence judgments significantly, p = .654. Findings indicate that the differences in intuitive coherence judgments across both lists are primarily driven by differences in thematic relations.

## References

Baumann, N., & Kuhl, J. (2002). Intuition, affect, and personality: Unconscious coherence judgments and self-regulation of negative affect. *Journal of Personality and Social Psychology*, *83*(5), 1213–1223. https://doi.org/10.1037/0022-3514.83.5.1213

Bolte, A., & Goschke, T. (2005). On the speed of intuition: Intuitive judgments of semantic coherence under different response deadlines. *Memory & Cognition*, *33*(7), 1248–1255. https://doi.org/10.3758/BF03193226

Bolte, A., Goschke, T., & Kuhl, J. (2003). Emotion and intuition. *Psychological Science*, *14*(5), 416–421. https://doi.org/10.1111/1467-9280.01456

Bowers, K. S., Regehr, G., Balthazard, C., & Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology*, *22*(1), 72–110. https://doi.org/10.1016/0010-0285(90)90004-N

Estes, Z., Golonka, S., & Jones, L. L. (2011). Thematic Thinking. In *Psychology of Learning and Motivation-Advances in Research and Theory. Advances in Research and Theory* (Vol. 54, pp. 249–294). Elsevier. https://doi.org/10.1016/B978-0-12-385527-5.00008-5

Fröhlich, J. K., & Högl, M. (2012). Thematic Ideation - Antecedents and Consequences of Individuals’ Thematic Similarity Recognition. *Creativity and Innovation Management*, *21*(4), 443–456. https://doi.org/10.1111/caim.12007

Kazén, M., Kuhl, J., & Quirin, M. (2015). Personality interacts with implicit affect to predict performance in analytic versus holistic processing. *Journal of Personality*, *83*(3), 251–261. https://doi.org/10.1111/jopy.12100

Kuhl, J., & Kazén, M. (2008). Motivation, affect, and hemispheric asymmetry: power versus affiliation. *Journal of Personality and Social Psychology*, *95*(2), 456–469. https://doi.org/10.1037/0022-3514.95.2.456

Lin, E .L., & Murphy, G. L. (2001). Thematic relations in adults’ concepts. *Journal of Experimental Psychology: General, 130*(1), 3-28. http://dx.doi.org/10.1037/0096-3445.130.1.3

Schönbrodt, F. D., & Hauser, J. (2015). *Does a mindfulness intervention performance in the remote associates test?*. LMU München.

Topolinski, S., & Strack, F. (2009). The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning. *Journal of Experimental Psychology. General*, *138*(1), 39–63. https://doi.org/10.1037/a0014678

Topolinski, S., & Strack, F. (2008). : Where There's a will- there's no intuition. The unintentional basis of semantic coherence judgments. *Journal of Memory and Language*, *58*(4), 1032–1048. https://doi.org/10.1016/j.jml.2008.01.002