Measuring Attention to positive subliminal cues via an emotional Stroop task that suggest a Pre-transcendent Fallacy held by Intimate Couples

Thomas A. Habib, Ph.D. and Christine A. Baser, R.N., Ph.D.

Abstract

An emotional Stroop task was utilized on 248 women to measure subliminal cues that were equated with the hypothesized pre-trans fallacy. This study was a 2 x 2 design that employed a dual task attention paradigm. There were two IV, 26 stimulus words presented subliminally, and two videos. The stimulus words had 3 levels ranging from romantic (adore), to non-romantic (helpful) to inanimate (door). All the words were lexically balanced and rated by women for category fit. The videos were of equal length and two levels, romantic and non-romantic. The romantic video involved a couple dancing the tango. The non-romantic involved a discussion on cloud formation. The DV is mean response time (RT) in an emotional Stroop task on a millisecond platform. It was hypothesized that both the romantic videos and the subliminally presented words would produce a slowing in RT compared to the other conditions. This was the first known use of an emotional Stroop test measuring powerful positive feelings pertaining to couples. A log-transforming data analyses was utilized. The hypothesized effect was ***not*** supported at any level. This may reflect that the hypothesis has no validity or that there is no emotional Stroop effect with positive words.

Introduction

The early romantic relationship is often characterized by idealized perceptions [1] and intense bodily experiences that easily obscure a more objective assessment of a potential mate. Although this experience is indeed wonderful and clung to with the hope of sustaining this intense feeling this investigation partially tests the proposition that the early romantic relationship plants a pre-transcendent fantasy in the hopes of many. There is evidence of an enduring expectation that love at this level is sustainable (Favero & Marciano, 2014) and there is documented impact of this depiction by popular culture (Hefner and Wilson, 2013). Consequently, Habib (2016) has suggested two widely held fears. First, the perception those other couples indeed get to live at this level of intensity and second the fear of a mistake in partner selection.

A pre-transcendent fallacy is an Integral Theory [2] concept. It is defined by Wilber (2000) as an inaccurately elevated and thus mistaken interpretation of an event or person… that is based upon enthusiasm and bodily sensations versus empathy and understanding that can only come with time and more fully developed knowledge. This tendency to idealize potential mates has been identified in the earliest stages of a romantic relationship and theorized by Habib (2016) to inhibit progression through the first three stages of a couple's development. It is hypothesized [3] (and will be tested in the future) that this intense pre-transcendent fantasy (PTF) is widely and unconsciously held by individuals in intimate

relationships consuming attentional resources, contributing to couple dissatisfaction, and otherwise inhibiting development.

This investigation is employing a dual task attention paradigm, the emotional Stroop task, to assess if there is prolonged attention to subliminal words associated with PTF unique to the intimate relationship. Numerous studies have suggested that when attention is drawn to emotionally valent words, subjects experience cognitive intrusion impairing performance on a Stroop task. This has been demonstrated, for example, when the emotional concerns involve anxiety including panic disorders, post-traumatic stress disorder (Williams et al., 1996, Matthews and Klug, 1993), sexual abuse (Freeman and Beck, 2000), depression (Joormann and Gotlib, 2007) or a threat to one's identity (Kaiser et al. 2006) . Specifically, there is a significant delay in mean response time (RT) for words subliminally presented that have emotional concern for the subjects presumably leaving less attentional resources for naming colors.

This is the first known use of an emotional Stroop test measuring powerful positive feelings pertaining to couples. In this study it is hypothesized that a threat is not responsible for a slowing in response time but rather a *preoccupation* with the persistent PTF widely held and specific to the intimate relationship. It is proposed that this study will find a similar delay due to the PTF that is consuming attentional resources and slowing RT. Finding empirical evidence of this attention consuming feeling is a small step toward validating this concept and the central role it plays in the Couples Line of Development (Habib, 2016).

Hypothesis

This study hypothesizes that women who are activated by a romantic video and instructions suggesting romance will respond to subliminal romantic stimulus words (Category 1) with a slower mean reaction time than positive trait words (about non-romantic people, (category 2) and a much slower RT than neutral words (category 3). There will be no differences between all 3 categories in the non-romantic video and instructions without romantic suggestions.

Participants

Women of all ages, single or coupled, over the age of eighteen to increase the chance they have had a romantic relationship, will be recruited to participate to increase the generalizability of potential findings.

Design

This study employs a 2x2 design. There are two IV, stimulus words presented subliminally and two videos. The stimulus words have 3 levels (category 1-3). The video has two levels, romantic and non-romantic. The DV is mean response time (RT) in an emotional Stroop task. . It is hypothesized that RT will be slowest for Category 1 words, then Catergory 2 words and fastest for Category 3 words.

Category 1
Attraction in a New Relationship
We have wonderful feelings when we first fall in love. Our enthusiasm about our new partner is peaking along with powerful chemistry. Please rate the following words that reflect a woman's experience in a brand new relationship, being sure to consider words associated with both feelings and bodily sensations.

If the word is a perfect fit for this category rate it 7. If the word doesn't fit this category at all rate it 1. Anything in between, rate 2 to 6.

Category 2
Positive Traits in People
In this category we are looking for words that describe a friend, coworker, neighbor, family or acquaintance where there is NO physical or romantic attraction.
Please rate the following words that reflect positive, but not romantic, feelings for people. If the word is a perfect fit for this category rate it 7. If the word doesn't fit this category at all rate it 1. Anything in between rate 2 to 6.

Category 3
Neutral Words
In this category we are looking for words without any friendship or romantic relationship content.   If the word is a perfect fit for this category rate it 7. If the word doesn't fit this category at all rate it 1. Anything in between, rate 2 to 6.

Generation of stimulus words for subliminal presentation

Three categories of stimulus words (IV with 3 levels) were developed and balanced for lexical equivalence including frequency of common use, word length, and known mean response time (RT) (Balota, Yap, et al. 2007)   as measured on a Stroop test. Previous emotional Stroop studies were critiqued by Larsen, Mercer, and Balota (2006) for failure to control for lexical attributes, concluding that reported differences in RT in many emotional Stroop studies may not solely been attributed to the IV.

A survey was completed online via *PsychData* platform. Subjects were recruited via e-mail or word of mouth and requested to complete the survey. A large list of stimulus words were rated on a Likert scale by thirty women for fit within each category. For each word they evaluated the extent to which each word best fits one of three categories. Category one words had to represent romantic feelings and bodily sensation associated with a new romantic relationship. There was ten Category One words: romantic, passionate, chemistry, hot, devoted, tender, adore, loving, muscular, and provider. Category two words were rated as to the extent to which each word represented non-romantic but positive attributes in a person whom a rater was not romantically attracted to like a coworker, neighbor, family member or acquaintance. There were six Category two words: reliable, helpful, consistent, dependable, honest and stable. Category three words were rated for having nothing to do with a description of a person. There were ten Category three words: bookshelf, automobile, chair, driveway, garage, pencil, television, window, doorway, and door. Mean rating, standard deviation, and percentage differences were plotted on a graph. The differences between category one and category two words were expectantly much smaller than the differences with category three words. This is why only six words were selected. It was concluded that if the results of this experiment were significant that the small differences between these two categories would only add to the importance of the hypothesized outcome. Also, in the final selection of category one words, the world compelling was ranked higher than provider. Provider was chosen by the researches who felt that it was ranked lower among women due to social desirable response factors.

Videos

The second IV with two levels was two videos selected from You-Tube. One video labelled romantic consisted of a heterosexual couple engaged in a tango dance. The other video was about identifying different types of clouds. The videos were 2 minutes and 48 seconds and 2 minutes and 33 seconds respectively in length.

Method & Data Collection

Data Collection

The experiment was completed online via *Millisecond* software and platform. Subjects were recruited as paid participants through *Amazon Mechanical Turk*, via e-mail, and word-of-mouth.

After consenting to participation, subjects will learn that they are completing a computer task that involves identifying the color of words (red, blue, green, yellow) by pressing corresponding keys.

This will be followed by viewing one of two videos before completing the Stroop task., either a romantic video or a non-romantic video. The instruction preceding the romantic video will suggest that this study is about maintaining the romantic intensity of an early relationship. The instructions preceding the non-romantic video will mention nothing about romantic relationships.

Each of the 26 priming words will be presented in black font for a total of 104 trials, in random order. Each trial will start with a standard foveal presentation of a fixation point for 1,000 ms. This point will be where the stimulus word will appear for 15 ms. Numerous research has shown 15ms exposure time to be consciously undetectable to the subjects. Each priming word will be immediately followed by a Stroop word and the participants are instructed to press the key corresponding to the correct color as quickly as possible. The consonants will remain on the screen until a key is pressed. Stroop words are a string of seven consonants, randomly ordered in red, blue, green, or yellow font.

For a randomly assigned half of the subjects in this study, the subliminal words will be removed and the subjects will be presented only the Stroop consonants with the same instructions to identify the color as quickly as possible. There will be no priming words in this portion of the study. The Stroop words are presented in random order in all four colors. Each trial will start with a standard foveal presentation of a fixation point for 1,000 ms., followed by a total of 104 trials of Stroop words.

Practice Trials

Each participant will be given 30 practice trials *prior* to data collection to practice seeing colors, hitting the correct key, and to familiarize themselves with the procedure and instructions. At the end of the practice, they are given a chance to review the procedure and practice again, or continue on.

Post Testing

 At the end of 126 testing trials the subjects were queried to rate the romantic strength of their respective video and whether or not they saw any real words. No subjects reported seeing anything but a string of consonants.

Random String of Consonants

26 random strings of seven consonants were generated using a random strings generator (randomstrings.org).

KVHRSAR
CBWQEUU
FIAFPPU
LJWBATV
CZEDQQI
HWICGSH
FHIQMVV
UOMTKBE
YMAVVWE
HICYCMV
PLPEUJA
MKEHNYV
SGPVOPW
REYZBHB
QMGKKKX
RKRGQIO
CBPBYBE
FYGLGKC
DFGADLO
OAVZFWG
JGEEZIH
URASHYQ
TMJQJPS
RCTTKGK
NRZKRKO
IOIYDWB

Selected Stimulus Words

A data analysis of the stimulus words was completed that included mean ratings of categories 1-3, mean differences of category 2 minus 1, category 3 minus 1, and category 3 minus 2 all of which was plotted. There were more statistically significant differences between Categories 1 and 3 and again between Categories 2 and 3 than there were between Categories 1 and 2. In most instances the scores with the greater mean differences were selected. The word "provider" was selected by the authors in place of the more highly ranked word "compelling" because we felt that this word was ranked much lower due to social desirability influences. All three categories were then lexically balanced for word length and

frequency of usage. This resulted in substituting the more highly ranked words (compelling, affectionate, tall) for the less highly ranked words (devoted, provider) to achieve this goal.

Category 1 (N=10)

Romantic, Passionate, Chemistry, Hot, Devoted, Tender, Adore, Loving, Muscular, Provider

Category 2 (N = 6)

Reliable, Helpful, Consistent, Dependable, Honest, Stable

Category 3 (N=10)

Bookshelf, Automobile, Chair, Driveway, Garage, Pencil, Television, Window, Doorway, Door

Instructions for Video 1: **Romantic Condition**

We are interested in the early romantic relationship and maintaining the feelings a couple experiences when they first fall in love. First you will watch a short video. After the video finishes you will be presented a series of letters. Your task is to identify the color you see from the four colors listed beneath them. The colors will be either red, blue, green or yellow. *Please do this as quickly as possible.* The speed of your reaction time is very important. Respond with your choice of color as quickly as possible, but also with accuracy. You will be given practice trials to help you complete this task. Make sure your speakers are on. Thank you for your help.

Instructions for Video 2: **Non-Romantic Condition**

This is a study about couples. First you will watch a short video. After the video you will be presented a series of letters. Your task is to identify the color you see from the four colors listed beneath them. The colors will be either red, blue, green or yellow. *Please do this as quickly as possible.* The speed of your reaction time is very important. Respond with your choice of color as quickly as possible, but also with accuracy. You will be given practice trials to help you complete this task. Make sure your speakers are on. Thank you for your help.

Results
The response of interest is the reaction time (latency), the time in msec from initial display of the letters until the cursor was depressed.

At the subject level, the design is a 2-by-2 design involving two Video Types [Romantic (R) and Non-Romantic (NR)] and two Word Types [Non-Prime (NP) and Prime (P)]. Subjects in the Non-Prime group were shown 26 random 7-character strings four times each, once under each color. Subjects in the Prime group were subliminally exposed to 26 actual English words four times each, once under each color. The 26 English words consisted of 10 that were classified as `romantic words', 6 that were classified as `people words' and 10 that were classified as `inanimate object words'. The actual words used are shown in Table 1 below. Over all subjects, there would be some interest in determining if there were significant deference's in

the mean response time in the four cells of the 2-by-2 design. For subjects in the Prime group, there is additional interest in knowing if the mean response time varied between the words in the 3 sub-groups (and whether this was effected by the type of video viewed). In particular, based on previous researchers' observations, the client hypothesized that women who had viewed the romantic video would become distracted by the subliminal romantic words and would respond more slowly to these, on average, than they did to the other two types of words.

Table 1
Words by Prime Group

#	Romantic	People	Inanimate
1	Romantic	Reliable	Bookshelf
2	Passionate	Helpful	Automobile
3	Chemistry	Consistent	Chair
4	Hot	Dependable	Driveway
5	Devoted	Honest	Garage
6	Tender	Stable	Pencil
7	Adore		Television
8	Loving		Window
9	Muscular		Doorway
10	Provider		Door

Data-Set and Inclusion Rules

Data were collected from 298 respondents, but not all of these could be used. In particular, results that were submitted by males or those of indeterminate gender were excluded, as were those from respondents under the age of 18. Several respondents became bored and didn't complete the exercise and a few were deleted for quality control issues. In the end, only 248 of the 298 respondents' data could be used. Table 2 below shows the reasons for exclusion.

Table 2
Sample Exclusion
Table Count & Reason

298 Original Sample
7 No Gender
30 Males
3 Underage Females
7 Incomplete Forms
1 Too Many Incorrect Answers
2 Response Times Too Variable
248 Final Sample Analyzed

One woman was excluded because only 30% of her responses (moving the cursor by color) were correct (all other respondents had at least 90% correct responses, most over 97%). Two other respondents were removed from the analysis despite having complete and mostly correct responses. These two respondents' variability in response times was much greater than the other 248 respondents, as discussed in the next section. This wasn't caused by one or two gross outliers but rather by variability throughout, so it was decided that it was better to eliminate these two subjects entirely. Of the 248 subjects included in the final analysis, the distribution by Video Type and Character Type is shown in Table 3 below. If the design

had been completely orthogonal, one would have expected 62 observations in each of the four interior cells of Table 3, but the distribution shown is within normal variation for random assignment of 248 subjects to the four cells.

Table 3 Distribution of Sample Subjects by Video Type and Word Type

	NonPrime	Prime	Total
Non-Romantic	59	57	116
Romantic	69	63	132
Total	128	120	248

Data-Cleaning

The response variable of primary interest is reaction time, or latency as it is called in the data-set. The utilized data-set (see selected.csv) has 248 * 104 = 25; 792 recorded latencies. This huge sample size allows one to obtain very tight confidence intervals for certain parameters, but also means that one must be quite vigilant about outliers that could easily crop in with so many observations. As others who have studied such data previously have noted, the distribution of latencies is very heavily right-skewed. This is not surprising, since there is a lower theoretical limit on how quickly a human can react to a stimulus, but no upper limit.

With respct to the lower reaction time limit, in Olympic sprint events, 100 msec is used as the critical reaction time below which a false start is declared. Other research indicates that about 170 msec is probably near the limit for reaction time to visual stimuli. A previous research article by Kaiser, Vick, and Major (2005) concerning reaction times with the same sort of cues used in the client's experiment indicated that these researchers used 300 msec as the lower bound for acceptable responses. We found a number of RT in the 200-300 msec which all had the correct color, so we tend to believe they are real. (i.e. the subjects wouldn't be able to get so many correct color identifications if they were responding before the subliminal message was completed.) For the very few reactions which were below 200 msec, we initially set then to 200 msec before deciding on further handling.

The upper bound is much more problematic. The afore-mentioned Kaiser article used 3000 msec (3 seconds) as the upper threshold for an acceptable observation. Even with this restriction, they found that the reaction times were strongly right-skewed and suggested, as most statisticians would, that the data be log-transformed before analysis. We concur with this assessment for the current data-set and will henceforth perform analyses with the log-transformed latencies. Although any log-scale could be used, we used natural-log (ln) and refer to our transformed variable as LNL, for `natural-log of latencies'. At this initial stage, we didn't exclude any latency for being `too large', although we made further adjustments as discussed next.

Upon log-transforming and performing simple analyses of the LNL values, we made several observations. Among these, as was to be expected, is that there is much variation in mean by subject. Less expected, but hardly surprising, is that there is much variation between subjects with respect to SD for the LNL observations. In addition, many subjects' distributions (as measured by examining the histograms of their 104 LNL values) are still slightly right-skewed even after the log-transformation of the reaction times. There are two subjects whose distri-

butions are especially problematic, with very skewed LNL distributions and much larger SDs than those of any other subjects, so these two subjects were dropped from the analysis, yielding the 248 subjects noted in the previous section.

The large observations in the upper tail of the LNL values can cause trouble with para- metric analyses such as ANOVA. While this can be remedied by using nonparametric methods, that seemed to be too much of a loss with respect to explanatory power. One possible solu- tion is to simply delete observations that are above a certain limit. Kaiser et al. mentioned using 3000 msec (or 8.006 in LNL scale), but that seemed to us to be an arbitrary cutoff. In addition, Kaiser et al. suggested deleting observations that are more than 3 SDs above or below a subject's mean and those on which a subject selected the wrong color. We disagree with these suggestions. First, if one wants to use any sort of _3 SD rule, one should first log-transform before applying the exclusion rule. Even then, however, simply deleting such observations as if they never happened definitely biases conclusions, unless one believes they are true outliers, such as might occur if the subject paused to eat or drink while performing the task. We really would like to keep all observations for a given subject, so that orthogonality, as much as possible, can be conserved.

Our solution was to first log-transform all reaction times. Then, for each subject, we ranked her log- reaction times from least to greatest and found the mean of the middle 84 (excluding the lowest 10 and the highest 10), as:

$$\bar{Y} = \frac{Y[11] + Y[12] + : : : + Y[93]] + Y[94]}{84}$$

where Y[j] represents the j-th ranked of the 104 LNL observations for the subject under investigation. For each subject, we also obtained a robust SD estimate from these middle 84 LNL values using the formula:

$$\alpha = \frac{Y[94] - Y[11]}{2 \times 1:2707}$$

:
We the retained all LNL observations which fell in the range given by:

$$[LB; UB] = \overline{\overline{Y}} \pm (2:60 \ x \ \alpha)$$

and set values whose LNL were below LB to the LB and LNL values that were above UB to UB. The file selected summary.xlsx summarizes this process for the 248 subjects, with the mean and SD of the raw and log-transformed values shown in columns G-H and I-J respectively. Columns K-L and M-N show the mean and SD of the raw and log-transformed values after making the boundary adjustments noted above. As noted at the bottom of thefile, these adjustments don't occur often, with 0.1% of the LNL observations subject to lower bound adjustment, 1.6% of the LNL observations subject to upper bound adjustment, and 98.3% unadjusted. In this way we are able to keep all 104 observations for each of the 248 subjects without being particularly affected by large outliers. This is a much more efficient procedure than simply deleting the high values, but doesn't allow one or two outliers to warp ANOVA results.

The second sheet (`sorted') of the selected summary.xlsx file sorts the subjects within each of the four Video*Prime cells of Table 3 by adjusted mean LNL value. Some individuals with especially large means or SDs within each group are highlighted in red, but we ultimately left all of these subjects in the final analyses. Subject-level results are shown in Table 4.

Table 4 Mean and (SD) of LNL at Subject Level by Video and Prime Type

	NonPrime	Prime	Total
Non-Romantic	6.556 (.200)	6.501 (.227)	6.529
Romantic	6.520 (.235)	6.501 (.240)	6.511
Total	6.537	6.501	6.520

The results in Table 4 are for the adjusted mean and SD of LNL over the approximately 60 subjects in each cell. The actual SD over the approximately 6; 400 LNL observations per cell is about twice as large, as the SD's shown in parentheses are between-sample means (over 104 latencies) within the same cell. The typical within-subject SD over the 104 responses is about 0.300. The overall pooled SD at subject-mean level of analysis is 0:227, so that with samples of the size used here, one would need a difference of at least .041 between cell means to conclude that any of the four cell means are significantly deferent from one another. The only significant deference at this level of analysis is that the 59 subjects in the upper left cell (Non-Romantic Video, Non-Prime Words) have slightly longer LNL than the other three cells; there is no statistically significant deference between any of the other three groups. The fact that we haven't adjusted for Color in these analyses is of no concern, since all subjects have 26 observations of each color among their 104 observations. While this preliminary analysis suggests that the hypothesis might be false, these results are hardly conclusive, since the three Prime sub-types (Romantic, People, Inanimate) are pooled together as `Prime', perhaps masking true differences between these. This topic is examined in more detail in the next two sections of the report.

Another data-cleaning aspect that we examined was whether or not the distribution of adjusted LNL values was significantly affected by the date on which the subject participated in the study. There were three methods utilized to solicit subjects: the pre-Mechanical Turk period (40 subjects participating from July 16 to September 14th), the Mechanical Turk MS period (36 subjects participating from September 15th to October 14th), and Other Mechanical Turk Period (172 subjects participating from October 15 or later). There was some thought that better quality data might have been collected in some periods than others. When we investigated this, we found a possibly significant difference between periods with respect to ineligible respondents (more occurred in the beginning), but none with respect to mean adjusted LNL values. Thus, we continued to use all 248 subjects' responses in performing the final analyses.

Main Effect Analyses
We used the same data-set as discussed above, but were looking more at the observation (LNL) level than the subject level, since we wanted to examine the effect of Color and of the three sub-classes of the Prime variable. For this part of the analysis, it is convenient to let `Prime0' refer to the non-English-word 7-letter words that were previously called Non-Prime and to let `Prime1', `Prime2', and `Prime3' refer to `Romantic', `People', and `Inanimate' words. At the

observation level, the sample sizes in each of the 8 cells are as shown in Table 5 below.

Table 5 Count of LNL Observations by Video and Prime Category

	Prime0	Prime1	Prime2	Prime3	Total
Non-Romantic	6136	2280	1368	2280	12064
Romantic	7176	2520	1512	2520	13738
Total	13312	4800	2880	4800	25792

The means and SDs within each of the eight cells are shown in Table 6 below. Note that the SDs within each cell here are about 0.40, as opposed to about 0.23 for the cells in Table 4. As mentioned previously, the fundamental reason for this discrepancy is that the SDs in Table 4 reflect variability in mean LNL level between the approximately 60 subjects within each cell. The SDs in the cells of Table 6 reflect not only the variability in means between subjects as well as the variability between the 104 individual LNL observations within each subject. One observation from the summary in Table 6 is that the Video Type seems to have almost no effect on mean time for any of the Primed (Prime1-Prime3) groups. Only for those who were given the non-Prime random letters did video appear to have any effect, with the Romantic video group performing slightly faster.

Table 6 Mean and (SD) of LNL at Obs. Level by Video and Prime Category

	Prime0	Prime1	Prime2	Prime3	Total
Non-Romantic	6.556 (.392)	6.491 (.381)	6.506 (.402)	6.509 (.399)	6.529
Romantic	6.520 (.394)	6.487 (.404)	6.515 (.402)	6.508 (.400)	6.511
Total	6.537	6.489	6.510	6.509	6.520

By way of comparison, the variability in mean LNL by Color is shown in Table 7 below. Each of these cell values is based on 25792=4 = 6448 observations. It seems that most subjects could react faster to Red and Yellow than they could to Blue or Green. This most likely has to do more with handedness than the effect of colors themselves, since the arrows which need to be pressed for Green or Blue responses are not as easy for most right-handed subjects to press as are those for Red and Yellow. The color effect appears to be somewhat more pronounced than the factors of main interest. Fortunately, the four colors are distributed evenly throughout the experiment, so it should have no effect on the variables of primary interest.

Table 7 Mean and (SD) of LNL at Obs. Level by Color

Color	Blue	Green	Red	Yellow	Total
Mean (SD)	6.544 (.413)	6.571 (.403)	6.489 (.385)	6.474 (.372)	6.520

If one constructs an additive main-effects model, neglecting the effect of subjects, one finds these deviations from the grand mean of 6.520

Video	[NR -0.009, R +0.009]
Primes	[P0 +0.017, P1 -0.031, P2 -0.010, P3 -0.011]
Color	[Blue +0.025, Green +0.052, Red -0.031, Yellow -0.046]

Final Analyses

This section contains the results of the Final Analyses. This was run in SAS 9.4 using PROC MIXED by Yan Du on March 6, 2017. The SAS code is reproduced in file Final SAS Code.docx. The output from this code is shown in three parts in file sasoutput.pdf. In each part, one or more Linear Mixed Models (LMM) was run on the data. In all of the linear models, the subject variance is estimated as .05371, while the random error's variance is estimated as .1054. This means that even if other factors such as Video Level, Prime Level, and Color are fixed, there is typical variability of the order of

$$p: \quad \sqrt{.05371} = .225$$

due to variation within the subjects. In addition, for a particular LNL observation within a subject, there is a typical error of the order of p:1054=.327 due to random error. These errors are both rather large with respect to the effects being estimated. The fact that there are about 60 subjects in each cell and 104 observations for each subject helps somewhat in differentiating effects, but the results are weak.

Analysis of Primes vs. Non-Primes

In part 1, the purpose is to compare the Non-Prime vs the Prime data, controlling for the fixed effect Video Type and the random effect Subject, both excluding and including the fixed effect of Color. As noted earlier, there is no significant effect due to 2-level Prime Type, Video Type, or their interaction. There is a signiffcant Color effect, with Blue and Green taking significantly longer (by about 0.08 in LNL scale) than Red and Yellow, but with no significant difference between Blue and Green nor between Red and Yellow.

Analysis of 4 Levels of Primes

In part 2, the purpose is to compare the 4 levels of the Prime data, controlling for the fixed effect Video Type and the random effect Subject, both excluding and including the fixed effect of Color. This is trickier than Part 1, since the levels 1, 2, and 3 of Prime occur within the same subjects. There seems to be something different (slower response) for the (Non-Romantic, Non-Primed) subjects, but that isn't of much interest. Within the Primed Group, there appears to be almost no effect due to Video Type, so in the last analysis, we considered the Primed Groups only.

Analysis of Prime Data Only

This is the analysis that most clearly shows the differences between the 3 levels (1-3) of the Prime variable. There appears to be no significant differences in the mean of LNL between levels 2 (People words) and 3 (Inanimate words), under Romantic or Non-Romantic videos. The 1-2 and 1-3 differences are negative, but not quite significant under the Non-Romantic video conditions. Under the Romantic Video conditions, these differences are even more negative and are both significant, with P-values of about 0.01 and 0.02 , respectively. Other than Color,

these are the only significant effects found in the final analyses.

Conclusion

One questions whether the significant results found in the last stage are real. The P-values (0.01 and 0.02) are small, but, then, we've conducted so many tests in this report that one might question whether a false positive has shown up just by chance. Even if the results are real, they are completely opposite of what the client expected. This would mean that women who are shown Romantic videos respond significantly more quickly when primed with `Romantic' words than they do when primed with `Person' or 'Inanimate' words. This is completely opposite of what the client hypothesized would happen.

References

Balota, D., Yap, M., Cortese, M., Hutchinson, K., Kessler, B., Loftis, B., Neely, J., Nelson, D., Simpson, G., and Treiman, R. (2007) The english lexicon project. *Behavior Research Methods* 39, (3), 445-459.

Freeman, J., & Beck, J. (2000). Cognitive interference for trauma cues in sexually abused adolescent girls with posttraumatic stress disorder. *Journal of Clinical Child Psychology*, 29(2), 245-256. doi:http://dx.doi.org/10.1207/S15374424jccp2902_10

Habib, T. (2016) A couples line of development. *The Integral Leadership Review*. August-November. Integral Publishers.

Hefner, V., & Wilson, B. (2013) From Love at First Sight to Soul Mate: The Influence of Romantic Ideals in Popular Films on Young People's Beliefs about Relationships *Journal of Communication Monographs* Volume 80, 2013 - Issue 2, 150-175. http://dx.doi.org/10.1080/03637751.2013.776697

Joormann, J., & Gotlib, I. H. (2007). Selective attention to emotional faces following recovery from depression. *Journal of Abnormal Psychology*, 116(1), 80-85. doi: http://dx.doi.org/10.1037/0021-843X.116.1.80

Kaiser, C., Vick, S., and Major, B. (2006) Prejudice expectations moderate preconscious attention to cues that are threatening to social identity. *Psychological Science*. Vol. 17, No 4, 332-338.

Larsen, Mercer, and Balota (2006) Lexical characteristics of words used in emotional Stroop experiments. *Emotion*. Vol. 6, No. 1, 62-72.

Mathews, A., & Klug, F. (1993). Emotionality and interference with color-naming in anxiety. Behaviour Research and Therapy, 31(1), 57-62. doi: http://dx.doi.org/10.1016/0005-7967(93)90043-T

Wilber, K. (2000) *Integral psychology* Boston, Shambhala Publications, Inc.

Williams, J., Mathews, A. and MacLeod, C. (1969) The emotional stroop task and psychopathology. *Psychological Bulletin*, 120, 3-24.