11-23-2018 10:00 AM

The Cognitive Psychology of Humour in Written Puns

James Boylan
The University of Western Ontario

Supervisor
Katz, Albert
The University of Western Ontario

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Cognition and Perception Commons, and the Cognitive Psychology Commons

Abstract

The primary purpose of this dissertation was to investigate how humour from written puns is produced. Prior models have emphasized that novel or surprising incongruities should be important for humour appreciation (Suls, 1972; Topolinski, 2014). In study 1, a new approach to operationalizing incongruity as semantic dissimilarity was developed and tested using Latent Semantic Analysis (Landauer, Foltz & Laham, 1998). "Latent semantic incongruity" was associated with humour ratings, but only for puns with low ratings of familiarity from a prior occasion or for those with a low level of aggressive content. Overall, there was also an unexpected strong positive association between familiarity and humour ratings. Study 2 demonstrated that humour ratings for puns decreases with repeated exposures. Changes in humour with repetition were dependent on latent semantic incongruity, the duration of time spent comprehending the pun and providing humour ratings, and on how humour was measured. Study 3 investigated whether "elaboration" on the two implied concepts in each pun was associated with humour (as predicted by Wyer & Collins, 1992). Elaboration quantity (the number of associated words that participants could comfortably list) and elaboration duration (the duration of time participants spent on the elaboration task) were positively associated with humour ratings, but only for familiar puns. Study 3 also found that participants who were assigned to focus on semantic dissimilarities found the puns to be more humorous, while participants who focused on semantic similarity produced a greater quantity of elaboration. In summary, fluent comprehension of incongruity was important for humour from unfamiliar puns, whereas elaboration on the implied concepts in puns was important to humour appreciation for puns that were familiar from a prior occasion.

Keywords

Humour appreciation, humor, semantic distance, incongruity-resolution, comprehension-elaboration, dissimilarity, similarity, familiarity, repetition, fluency

Co-Authorship Statement

All projects in this dissertation were conducted under the supervision of Dr. Albert Katz. The written material in this thesis is my own work. Albert Katz provided input and assistance with editing and revision of the content. Data from study 1 was initially collected as a part of a collaborative project with Dr. Rod Martin. Dr. Martin provided preliminary input and edits for study 1 and also developed and hosted the internet-based survey platform that was used to collect data for all reported studies.

Acknowledgments

I would like to express my sincere and enthusiastic gratitude to my supervisor, Dr. Albert Katz. His generous support, reliability, encouragement, and commitment to his students had a tremendous impact on my ability to succeed. I feel lucky to have had the opportunity to work with him. Albert has been consistently kind, positive, and understanding. Moving forward, I hope that I will be able to live and work following his example.

I would like to thank my dissertation committee including Dr. Rod Martin for both his technical expertise, input, and assistance in shaping the early phases of this project. Thanks also to Dr. Nick Kuiper and Dr. Debra Jared for their input as committee members and for their thorough and thoughtful review of this dissertation.

I would also like to thank my lab mates and colleagues in the cognitive psychology graduate program at Western University. My thanks to Hamad Alazary, Nick Reid, Jason Perry, Alex Taikh, and everyone else from our program. I am grateful for the friendly and positive work culture we developed together. Thank you for your insights, comments, and support.

On a personal level, I would like to thank my sister, Amy Boylan and longtime friends who are like family; including Rick Coupland, Sarah Cazakoff, Scott Davison, and Demetrius Michael. Your encouragement and support really kept me going.

I would like to thank my mother, Susan Boylan, whose strength, patience and grace taught me not only to survive, but to thrive despite life's adversities. You are a reliable role model who I am lucky to have in my life. I would also like to express my gratitude to my late father, John Boylan, for his unconditional love, reliability and consistency. Thank you for teaching me to stick with things even when they are hard, to love learning new things, to be there for people you care about, and to love life.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

1 General Introduction

The primary purpose of this dissertation was to investigate how humour from written puns is appreciated. Predictions from empirical models of humour appreciation were operationalized and tested using techniques from cognitive psychology. It has been argued that humour helps us enjoy working though complex information and generating novel insight (Amir, Biederman, Wang & Xu, 2013). Humour is important and ubiquitous to everyday human life but there has been relatively few empirical studies of longstanding theories and assumptions regarding how humour is processed. Across three studies, predictions regarding the function of incongruity, familiarity and repetition of humorous stimuli, aggressive content, fluency of processing, and elaboration for humour appreciation were empirically tested.

There is a long-standing consensus that incongruity is important to humour appreciation (Suls, 1972; Schultz, 1972; Wyer & Collins, 1992; Hillson & Martin, 1994; Attardo, 1997). However, there has been considerable debate regarding how incongruity in humour should be defined and operationalized for empirical study (for a discussion see Ritchie, 1999; 2004; 2009; Forabosco, 1992). There are potentially many sources of humorous incongruity, such as: social incongruity (from atypical or inappropriate social situations), surprise or violation of expectations, aggression, tension relief, and from events with low typicality (such as nonsensical silliness). The current research operationalizes and examines incongruity in humour as semantic dissimilarity between two implied concepts in written puns.

Incongruity in humour has been challenging to operationalize for experimental study, with some arguing that the greatest ongoing challenge to the development of a comprehensive theory of humour has been that incongruity has not yet been clearly, or consistently, defined. In fact, it has been said that few formal models provide a definition of incongruity that is more precise, detailed or formal than a dictionary definition of the term (Ritchie, 1999; 2004; Forabosco, 1992). Study 1 developed a new measure of

semantic incongruity, using Latent Semantic Analysis (LSA). Latent Semantic Analysis is a mathematical measure of the semantic similarity in meaning between referenced words or text passages (Landauer, Foltz & Laham, 1998). In the currently reported studies, items with low semantic *similarity* between concepts according to estimates from LSA are considered to have high latent semantic *incongruity*.

The core goal of study 1 was to investigate the association between semantic incongruity and humour in written puns using a newly developed measure of latent semantic incongruity. In study 2, latent semantic incongruity estimates were applied to investigate the "resolution" of semantic incongruity effects in puns that are presented multiple times for participant assessment (as hypothesized by Suls, 1972). In study 3, the latent semantic incongruity measure was further applied to investigate specific predictions from the comprehension-elaboration model (Wyer & Collins, 1992). Semantic incongruity was investigated as an indicator of the potential of a pun to generate novel elaboration on associated concepts.

1.1 Written Puns

The function of semantic incongruity in humour appreciation was examined using puns as experimental items. Thus, prior to discussing empirical findings relevant to incongruity and humour, the relevant characteristics of puns will be described.

1.1.1 Definition

Puns are short sentences that each contain a key polysemous word that is orthographically (based on identical or similar spelling) or phonologically (based on identical or similar pronunciation) ambiguous in a way that creates semantic incongruity given prior sentence context (for example as seen in the pun, "Coming up with cheese puns should be a *bries*"; see Duchacek, 1970). Puns are distinct from ambiguous yet non-punning sentences because they preserve two incompatible meanings at the end of a sentence (Attardo, 1994). The orthographic or phonological ambiguity between multiple implied meanings creates an incongruous semantic contrast: in this case between the "bries: cheese" concept and the implied "breeze: easy" concepts. Both meanings can potentially make sense in the semantic context of the sentence and the alternate

interpretations play off of each other. Thus, according to the operationalization of incongruity as semantic dissimilarity in the present studies, high levels of semantic incongruity (indicating low levels of latent semantic similarity in LSA) between the concept "bries: cheese" and the concept "breeze: easy" should be predictive of the extent to which each pun is humorous.

1.1.2   Polysemy Mechanisms in Puns

Duchacek's (1970) taxonomy (as cited by Attardo, 1994), categorized puns according to the linguistic mechanism of the key polysemous word in each item. Duchacek identified several main categories: puns based on homophones (such as "whole" vs. "hole"; words with identical phonemic representations but different spellings), homographs (such as "bore: to drill" and "bore: dull or uninteresting"; words with identical orthographic representations; that are spelled the same), and rhyme based puns (referred to as "paronyms; such as "braid" vs. "grade" or "mother" vs. "another"; words with similar but not identical orthographic and phonemic representations, based on rhymes or morphological combinations. Puns can be created in other ways (for example using antonymy or syntactic reversals) but these alternate mechanisms (as discussed in Attardo, 1994) are not as clearly defined, frequent, or distinct as pun mechanism categories. The studies reported in this dissertation thus investigated humour appreciation using puns based on homophones, homographs, and rhymes as distinct categories of puns.

In the present studies, all homograph puns were based on words with both identical spelling and pronunciations but with different implied meanings, such as in the pun "a bad shoemaker's assistant was given the ***boot***". This pun was based on the implied concept "boot: the foot covering" and the concept "boot: to dismiss from employment". Although homographs can have different pronunciations (as in the case of "bass: fish" and "bass: guitar"), the current study examined only homographs with both identical spelling and pronunciation. The homophone based puns in the present study used words that have identical pronunciation, but different spellings and meanings such as in the pun "A baker stopped making donuts after he got tired of the ***hole*** thing". This pun creates a contrast between the implied meanings of the words "***hole***: an opening into or through a pastry" and "***whole***: entire". Rhyme based puns represent a more diffuse pun category in

which two meanings are contrasted based on words with only a general similarity of sound and overlap in spelling. For example, the rhyme based pun "A dentist pulled out my tooth without meaning to, it was ***accidental***" creates a semantic contrast based on similar sounds and the overlap in spelling between the implied words "***dental***" and "***accident***".

1.1.3   Ideal Experimental Items

Puns are ideal as experimental items for the study of semantic incongruity because they tend to be of a similar short length, a similar low level of complexity, they are frequently experienced in everyday life and they are among the simplest examples of humour. More complex forms of humour such as multi-sentence jokes, cartoons, or comedic routines, are more likely to involve multiple sources of incongruity (that is, uncontrolled sources of incongruity other than a simple semantic contrast) and risk for other potentially confounding variables, such as personal biases and preferences that can be more difficult to isolate and quantify. The potential downside of using real world examples of humour is that people may already be familiar with a pun from a prior occasion. However, familiarity effects can be controlled, or even investigated as a factor of interest (as it was in study 1), by simply asking participants to self-rate the extent to which they are already familiar with each pun from a prior occasion.[1]

1.1.4   Puns are Ubiquitous

Although puns have been described as the lowest and least enjoyable form of humour (Dryden, 1672), they appear to be historically and cross-culturally ubiquitous. For example, Shakespeare (1564-1616), frequently used verbal puns in his work; in Romeo and Juliet (1597), Shakespeare played on the similar sound of two words with distinct meanings when he had Romeo say to Mercutio: "you have dancing shoes with nimble soles, I have a soul of lead." (Shakespeare, Romeo and Juliet. 1.4.14-15). There are orthographic Chinese puns based on similarities in character shape between words and

---

[1] Familiarity with a pun from a prior occasion, as referred to in the reported studies, indicates that participants have had an episodic experience with a pun from everyday life.

puns in American Sign Language based on ambiguities of articulation in hand location, hand shape, movement and orientation. For example, such as when a person makes the "S" symbol adjacent to his or her ears to make "Sears" (Attardo, 1994). In the modern world puns appear most everywhere: in advertising, news, prose and poetry, and often even in journal article titles. For example, see: "Gorillas in our *midst*: Sustained inattentional blindness for dynamic events" by Simmons and Chabris (1999; playing on the title of the book and movie "Gorillas in the mist"), or "Smells like *clean* spirit: nonconscious effects of scent on cognition and behavior" by Holland, Hendriks and Aarts (2005; a play on the title of the Nirvana song "Smells like teen spirit").

1.2   Incongruity Based Theories of Humour

1.2.1   Incongruity-Resolution Theory

Incongruity-resolution theory (Suls, 1972; see also Schultz, 1972) holds that incongruity is only humorous if it can be adequately explained (and therefore resolved). Suls (1972) argued that recognition and processing of humorous incongruity can be seen as an exercise in problem solving: people form an initial prediction of a situation given available context. If an initial prediction is violated it causes surprise and requires re-interpretation. If the incongruity from the violated expectations can be explained, it will be enjoyed, if it can't be explained it will end in non-humorous puzzlement. Suls predicted that humour should depend on: the level of incongruity in a joke, the complexity of the problem-solving task (ideally neither too hard nor too easy to understand; a non-linear inverted-u shaped function), the time taken to solve the problem, and the salience of the joke's content. Suls hypothesized that emotional content (such as aggressive themes) should serve to facilitate incongruity processing by making the content necessary to resolve the incongruity of an item more salient (Goldstein, Suls & Anthony, 1972). Incongruity-resolution remains a popular explanatory framework that has been frequently used or adapted, to some extent, in more recent models. More recent models make slightly different predictions regarding the relative role and importance of incongruity and resolution processes, but most posit at least a minimal role for resolution as the reduction of incongruity or explanation of humorous stimuli (for a review see Ritchie, 1999, 2004; Forabosco, 1992; Martin, 2003).

Predictions from the incongruity-resolution model hold that an intermediate level of difficulty in comprehension of incongruity should be optimal for humour. The strongest support for this hypothesis has come from developmental studies. Zigler, Levine and Gould (1966) investigated humour appreciation in children in second to fifth grades of elementary school. Humour appreciation for cartoons increased with the extent to which the students could understand the cartoons up to the fourth grade, but then decreased despite increased comprehension at the fifth grade: the cartoon stimuli seemed too simple for the fifth graders to enjoy. A subsequent study with a greater range in cartoon difficulty demonstrated that children in each grade found cartoons with an intermediate level of difficulty to be the most humorous (Zigler, Levine & Gould, 1967; see also McGhee, 1976).

There is a body of empirical studies that have identified incongruity-resolution as important to humour appreciation. For example, Ruch (Ruch, 1992; Ruch 1981, 1984, Ruch & Hehl, 1998) asked participants to rate a series of jokes and cartoons and conducted a factor analysis on positive and negative participant assessments (such as funniness or averseness). They identified three meaningful factors that could be used to categorize potentially humorous stimuli: humour could be produced from incongruity-resolution, from sexual themes, and from nonsense. Nonsense based items were those with incongruity but that had no clear resolution. Despite the fact that their studies contained participants with a wide range of backgrounds, and jokes with a wide range of themes (including items with aggressive content), only these three categories consistently stood out as distinct latent factors.

Incongruity-resolution was developed with joke and cartoon stimuli in mind. The model should however also apply to humour from puns. It could be speculated that people would analogously form an initial interpretation of a sentence that is violated by a recognition that an alternate interpretation of a sentence is possible. For example, in the pun "I didn't want to buy leather shoes, but eventually I was suede", the incongruous violation of expectations might occur when realizing that "suede" doesn't grammatically fit at the end of the sentence. The incongruity might be sufficiently explained upon recognizing that an item is a pun and that the alternate homophonic "swayed" concept

would also fit in the sentence. People could realize that a pun is intended and the two concepts could be further compared or contrasted to evaluate how well each concept fits in the sentence.

The hypothesis evaluated in the current research is whether semantic dissimilarity between these two possible concepts is predictive of humour. The assumption is that the violation of expectations from more semantically incongruous concepts would be more appropriately challenging to explain (neither too difficult nor too easy to understand) and therefore more humorous. People may also find puns based on semantically dissimilar concepts to be more humorous because the contrast may be seen as being more clever, unexpected or novel; presenting a more interesting or surprising challenge for comprehension.

Critics of the incongruity-resolution model have argued that resolution is not strictly necessary for something to be humorous. O'Shannon (2012) argued that if humour depends on incongruity, then resolution (as the explanation and reduction of incongruity) can only serve to decrease perceived humour. He argued that if your goal is to be funny, resolution should therefore be avoided as much as possible. In many cases resolution of incongruity is not logically or practically possible (cases that would correspond to Ruch's nonsense humour factor; for a review see Ruch, 1992). In some cases, attempted explanation of incongruity can actually introduce further incongruities (for a review see Forabosco, 1992, 2008; Attardo & Raskin, 1991, Ritchie, 2004). Pien and Rothbart (1977) discussed incongruity without clearly possible resolution using the example nonsensical joke "Why did the elephant sit on the marshmallow? Because he didn't want to fall into the hot chocolate". Here, the initial incongruity is technically explained, but not in a way that makes complete logical sense given external reality. Here the punch line to the joke, traditionally seen as a source of resolution, actually introduced new sources of incongruity (Pien & Rothbart, 1977; Ritchie, 2004). In response to these arguments, Suls (1983) argued that the incongruity-resolution model does not strictly require that explanations must be complete or perfect. However, superior resolution (clear comprehension and explanation with fewer loose ends) should provide more humour than less clear or complete resolution. In the case of non-sense humour, Suls argued that a

willing temporary suspension of disbelief could allow for sufficient resolution of nonsense humour.

1.2.2   Process Fluency Account

The process fluency account of humour builds and expands on incongruity-resolution based predictions (Topolinski, 2014). Whereas Suls (1972) predicted that humour comes from feelings of achievement as if it were a satisfying reward for completing a challenging comprehension puzzle, Topolinski (2014) predicted that humour is produced when processing challenges can be completed both quickly and easily (for consistent evidence see also Goldstein, 1970a). In support of the fluency account, Topolinski (2014; study 1) found that priming participants with material from the punch line of a joke prior to hearing it significantly increased funniness ratings (when primed 15-minutes or 1-minute prior to exposure to a joke; but not when primed immediately prior to the joke). According to the fluency account, priming content from the punchline of joke stimuli increased funniness ratings by facilitating the ease of incongruity comprehension. In a subsequent study, Topolinski found that jokes written in an easy to read font were rated as being funnier than jokes written in a harder to read font.

From the perspective of the fluency account, incongruity should be positively associated with humour only when the incongruity can be rapidly and easily processed. Topolinski (2014) used traditional "joke" stimuli with an incongruous build up and an explanatory punchline. It could be speculated that the fluency account should hold that puns should be more humorous if they can be processed more rapidly or if they are easier to understand. Semantic incongruity should be associated with humour if it can facilitate processing speed.

1.2.3   Comprehension-Elaboration Theory

The comprehension-elaboration model also builds and expands upon incongruity-resolution based predictions (Wyer & Collins, 1992). A frequent argument against incongruity-resolution has been that it places excessive emphasis on mere-comprehension (a valid challenge to incongruity-resolution, as acknowledged by Suls, 1983). Wyer and Collins (1992) predicted that humour appreciation can come from both the challenge of

explaining humorous incongruities (the degree of comprehension challenge) and on elaboration on the humorous aspects of the stimuli after comprehension. They argued that comprehension of incongruous stimuli should demonstrate a non-linear inverted-u shaped association with humour appreciation; that is, the incongruity should be neither too easy nor too difficult to comprehend (a similar argument was made by Suls, 1972; see also Berlyne, 1960). Beyond comprehension, Wyer and Collins (1992) proposed that humour can also come from an "elaboration" process that creates new post-comprehension connections between mental schemata. It was argued that relevant schemata that are activated by a humorous item should play back and forth off of each other activating further concepts and new mental images. Humour from the elaboration process was hypothesized to require the conscious generation of additional inferences about features that were not captured by an initial encoding. It was hypothesized that there should be a linear association between the duration of time that participants are *willing* to spend elaborating on the content of a joke after comprehension and feelings of humour. Wyer and Collins (1992) also argued that some form of re-interpretation and a minimal amount of diminishment (in terms of importance or value in comparison to a first interpretation) should be necessary for humour appreciation.

From the perspective of the comprehension-elaboration model, semantically incongruous puns should be humorous if they present a moderate level of challenge for comprehension or if semantic incongruity facilitates elaboration. Wyer and Collins (1992) speculated that incongruity in puns should be extremely easy to understand and so it should be unlikely that humour from these items should be produce by differences in the challenge of comprehension. They argued that humour from puns is more likely to come from the quality and quantity of elaboration on the implications of a pun. For example, in the pun "I used to be a Velcro salesman but I couldn't stick with it", it should be relatively easy to recognize that stick (to persist) and stick (as an adhesion between substances using Velcro) are possible interpretations of the homograph. Humour should be more likely to come from additional elaborations, such as reflections that someone may have actually failed at their job, one might wonder how hard it is to sell Velcro, or by imagining how an unsuccessful Velcro sales call might progress. Based on comprehension-elaboration predictions, it could be speculated that semantically

incongruous puns should be more humorous if they present a more appropriate challenge for comprehension or if the contrasted concepts can bring in a greater quantity of related content that can be used for elaboration.

1.3   Potential Moderators of Semantic Incongruity

1.3.1   Familiarity

Familiarity with stimuli from a prior occasion is an important potential moderator of the association between latent semantic incongruity and humour. Predictions from the incongruity-resolution model (Suls, 1972) emphasize the importance of surprise and novelty. The implication of this assumption is that jokes should not be funny, or at least as funny, when heard multiple times. Consistent with this argument, Herzog and his colleagues asked participants to rate a series of jokes and found that self-rated surprise was significantly positively correlated with funniness ratings (Herzog & Bush, 1994; Herzog & Karafa, 1998). Herzog et al. further argued that humour from feelings of surprise (or shock) can help explain people's enjoyment of "sick" macabre jokes (Herzog & Bush, 1994; Herzog & Karafa, 1998). However, people can clearly find a favorite joke, TV show, or comedy routine humorous multiple times. For example, Gavanski (1986) found that participant mirth (measured according to experimenter judgment of participant laughing or smiling responses), but not participant cognitive appraisal (overall assessment of each cartoon), significantly (but slowly) decreased with repeated exposures. The decreases in mirth were less than would be expected according to a strong interpretation of the incongruity-resolution model. After 5 repetitions of a cartoon, the average mirth ratings were still greater than 2 (on a Likert-type scale from 0-5). Further, Kenny (1955) found that the degree of predictability of punch lines in jokes was actually positively correlated with perceived humour, suggesting that participants may, on average, actually prefer jokes that are unsurprising.

The incongruity-resolution model (Suls, 1972) would have difficulty accounting for repetition effects: people can find a familiar humorous item funny multiple times; past the point that incongruities are no longer surprising or satisfying to comprehend. In fact, people can potentially find predictable content more humorous with repetition, as in a

running gag (for supportive evidence see Kenny, 1955; Schick, McGlynn & Woolam, 1972; Suls, 1975). The preservation or increase of humour with repetition is difficult to accommodate within a strong interpretation of the incongruity-resolution model. Suls (1972) argued that forgetting could potentially account for repetition effects. Suls argued that highly humorous items should become less humorous with repetition because they should be more accurately remembered. Items with low to moderate effectiveness should remain humorous on repetition, as they are more likely to be forgotten and again require resolution. Suls hypothesized that repetition effects could alternatively also be caused by mere-exposure effects (a process by which the positive assessment of stimuli will generally increase as people become familiar with stimuli; Zajonc, 1968), or by re-interpreting humorous stimuli (discovering additional incongruities and therefore also new resolutions of an item).

Suls' (1972; 1975) incongruity-resolution based hypothesis that mere-exposure effects can account for humour on repetition is consistent with the fluency account (Topolinski, 2014) given that increased familiarity with repeated exposures to stimuli is also a strong positive associate of verbal fluency (Zajonc, 1968). The fluency account would hold that repeated stimuli should remain humorous (or potentially even increase in humour) as stimuli become more familiar and therefore also easier to process. Wyer and Collins (1992) argued that mere-exposure should be insufficient to account for humour on repetition. Mere-exposure should operate equally on all items; however, some stimuli are more likely to retain humour on repetition than others. The comprehension-elaboration model (Wyer & Collins, 1992) makes an allowance for familiar items to remain humorous as long as people are still willing and able to elaborate: to generate novel insights in association with repeated stimuli.

1.3.2   Aggression

Aggressive content is another potentially important moderator of the association between latent semantic incongruity and humour. To minimize ethical concerns, only puns with low to moderate levels of aggressive content were selected as experimental items for the currently reported studies. Participant self-rated perception of the level of aggression in the selected puns was measured as a moderator variable of interest. It has been argued

that aggression is necessary and, varying from strong to weak instantiations of the argument, sufficient in itself for humour. Zillmann, Bryant and Cantor (1976; see also Martin, 2003) traced this logic historically back to Thomas Hobbes who saw laughter as being inherently based on feelings of superiority coming from the recognition of inferiority in another person. According to superiority theory, the pleasurable feelings accompanying humour are (on some level) essentially feelings of superiority over others arising from, a degree of ridicule, disparagement, victory and/or loss. From this perspective, humour should be seen as being inherently aggressive. Gruner (1997) argued that puns evolved out of a history of verbal jousting, in which people attempt to show off their verbal wit, skillful fluency, or cleverness. Given that puns do not necessarily have an explicit "victim" or active disparagement, superiority theory would view the interpreter of a pun as being implicitly forced into the role of a less witty "loser" in the interaction. This may help explain why some people may groan with displeasure in response to a pun: that is, making a pun may be an inherently aggressive act (Gruner, 1997). The claims of superiority theory are, arguably, overly broad and challenging to falsify (if all humour is, on some level, a form of aggression). However, at a minimum, aggressive content is a well-known associate of humour. Zillmann and his colleagues argued that a moderate level of aggression is optimal for humour (Zillmann & Bryant, 1974; Zillmann, Bryant & Cantor, 1974). In support of this argument, Zillmann, et al. (1974) found that participants prefer cartoons with a mild level of aggression directed towards political candidates, as opposed to cartoons with an excessive level of aggressive content or brutality (this effect was only significant for cartoons against candidates they did not support). According to this prediction from Zillmann et al., for the puns with a low to moderate level of aggression in the currently reported studies, aggressive content should be positively associated with humour.

Prior theorists have argued that aggression and incongruity should interact productively to enhance humour appreciation. Koestler (1964) argued that humorous incongruity ("bisociation") requires at least some aggression for it to be considered humorous. Misattribution theory, as proposed by Zillmann and Bryant (1980), holds that the purpose of incongruity in humorous material is to make aggressive content seem more socially permissible and therefore more humorous. Incongruity was hypothesized to help sanitize

aggression making it seem less offensive and therefore more humorous. It allows people to misattribute their enjoyment to the incongruities, when it may actually have been produced by the aggressive content. Suls' (Suls, 1977; Goldstein, Suls & Anthony, 1972) incongruity-resolution based salience hypothesis made a similar but causally distinct prediction that incongruity and aggression should interact productively for humour. Suls predicted that the purpose of aggressive content and other variables in humorous stimuli that are extraneous to an item's core incongruities (such as sexual or offensive content), is to increase humour by facilitating resolution (Suls, 1977; Goldstein, Suls & Anthony, 1972). In accordance with the comprehension-elaboration model, Wyer and Collins (1992) predicted that aggression should enhance OR inhibit humour depending on participant preference. If an interpreter is offended, or distracted by the motives of a speaker, they might spend less time elaborating on the incongruity at play and therefore they should find a joke less funny. If aggressive content grabs attention, and encourages elaboration, then aggression should enhance humour.

1.4   Investigation of Puns in Cognitive Psychology

Although their finding was not replicated by Jared and Bainbridge (2017), Kao, Levy and Goodman (2015) investigated semantic ambiguity and semantic distinctiveness as predictors of humour. Their measure of semantic ambiguity assessed the extent to which sentence context of a pun equally supports both meanings of a pun (which they argued may represent incongruity). Their measure of distinctiveness assessed the extent to which the alternate implied meanings are supported by subsets of words in the sentence of the pun (which was argued to represent the extent to which incongruities can be resolved). They found that semantic ambiguity could distinguish puns from control sentences, but only semantic distinctiveness was associated with participant humour ratings. In their computational model, Kao et al. (2015) represented the concepts in homophone base puns according to the appropriate spelling of each sense. In the example pun "The magician got so mad he pulled his hare out.", they compared the meaning of "hare" the rabbit with "hair" which grows out of follicles on the head. They acknowledged that this was only an approximation that captures the "gist" of the implied meanings in a pun.

Both the dominant and subordinate implied meanings in written puns appear to be activated and involved when processing written puns. This is relevant to the reported studies in this dissertation because a measure of semantic incongruity was prepared from the semantic dissimilarity between the two implied concepts in written puns. McHugh and Buchanan (2016) investigated semantic ambiguity processing in homograph based puns. Although they did not investigate funniness ratings in relation to their variables of interest, they demonstrated that the semantic relatedness of the alternate implied meanings in homograph based puns were predictive of priming effects in a lexical decision task. Their results indicated that both the dominant and subordinate possible meanings of homographs in pun sentences are activated. A similar result was obtained by Jared and Bainbridge (2017) for homophone based puns. Homophones with more frequent subordinate meanings had shorter gaze durations and shorter total reading times. Jared and Bainbridge interpreted this as indicating that the subordinate implied meaning of the homophone was activated through shared phonology.

Jared and Bainbridge (2017) obtained several findings relevant to the cognitive psychology of humour in written puns. They conducted an eye tracking study which compared reading times for the homophones in puns (such as "the butcher was very glad we could meat up") with reading times for the same homophones in control sentences which support only one possible interpretation of the sentence (the butcher was very glad to chop meat up for stew). First fixation durations on the homophones were longer in puns than for the control sentences but total sentence reading times were not different between puns and control sentences. They interpreted this as indicating that their participants immediately noticed that the homophone was incongruous given prior sentence context but the incongruity could be rapidly resolved. Jared and Bainbridge's (2017) finding differs from prior work on homograph base puns by Sheridan, Reingold, and Daneman (2009) who found that first fixation durations on homographs were shorter for puns than for control sentences.

Jared and Bainbridge (2017) found that participant funniness ratings were associated with longer gaze durations on the homophones (indicating recognition of incongruity) and shorter total fixation durations (indicating that the incongruity could be quickly resolved).

The strongest predictor of humour was the semantic similarity of the presented version of the homophone with a critical word from the prior context of a pun (e.g. between "meat" and "butcher" in the aforementioned example). They speculated that the puns may be more clever or unexpected when the context suggests a strongly related interpretation of a homophone but an alternate meaning is presented instead. There was also a trending (but not statistically significant) association between participant funniness ratings and phonological decoding skills ($r = .26$, $p < .10$), and there was no significant effect of homophone frequency on humour ratings.

1.4.1   Operationalization of Incongruity as Semantic Dissimilarity

Trick and Katz (1986; see also Tourangeau & Sternberg, 1981) used the "domains interaction" approach to demonstrate that metaphors with greater semantic dissimilarity between concepts are more accurately understood and are more appreciated. Participants were asked to rate the qualities of 27 individual concepts (e.g., people, animals, etc.) used to form metaphors according to relevant semantic dimensions (e.g., human likeness, classiness). The 27 concepts were used to form 306 metaphors of the form "A is the B of A's domain (for example, as in the metaphor: "The Concorde is the mosquito among aircrafts"). The metaphors were each rated for their level of comprehensibility, aptness, ease of interpretation, and the extent to which participants liked each metaphor. Trick and Katz conducted a factor analysis on semantic dimensions used to rate each of the concepts. Two factors were found to be domain distinguishing (made up of variables on which the metaphor concepts could be dissimilar to each other) and two factors were found to be domain-insensitive (made of variables on which metaphor concepts could be similar to each other). For example, *George Bush* and a *car* can have similar levels of classiness but would have dissimilar ratings for the extent to which they are humanlike. The semantic relatedness between the two concepts in each metaphor according to semantic factor loadings was calculated using a Euclidian distance formula. They found that metaphors with high loadings on the domain distinguishing factors (thus having greater dissimilarity between the contrasted concepts) and low loadings on the domain insensitive factors (thus also having few ways in which the concepts could be similar to each other) were easier to understand and were considered to be more apt as metaphors.

Hillson and Martin (1994) replicated the Trick and Katz (1986) domains-interaction study design using the same metaphor item set, but referred to the metaphor items instead as jokes, and examined participant humour ratings for the items in relation to semantic relatedness. The domain distinguishing factors (taken as a measure of semantic dissimilarity or incongruity) were found to be significantly correlated with humour, but domain insensitive factors (taken as a measure of resolution or semantic similarity; meaningful ways that the concepts in each metaphor can be similar to each other that can be used to resolve incongruities) were not found to be significantly associated with participant humour ratings. They also obtained a significant interaction wherein the funniest metaphors had high loadings on both incongruity (dissimilarity based) and resolution (similarity based) factors. Hillson and Martin (1994) demonstrated that a measure of semantic relatedness can serve as an effective operationalization of incongruity. This approach allowed incongruity to be quantified, measured, and compared against participant humour ratings for artificial metaphor stimuli. The current dissertation will build and expand on this approach to further study the function of incongruity and resolution in the context of puns as real-world examples of humour (as opposed to the artificial lab-created metaphors used in prior studies).

1.4.2   Overview

The three reported studies examined semantic dissimilarity as a predictor of participant humour appreciation. In study 1, a new measure of latent semantic incongruity was developed using latent semantic analysis in written pun stimuli based on homophones, homographs and rhymes (Landauer, Foltz & Laham, 1998). This first study examined the new "latent semantic incongruity" measure as a predictor of humour appreciation in relation to both participant ratings of familiarity with a pun from a prior occasion and to the presence of aggressive content.

In study 2, latent semantic incongruity was further examined as a predictor of changes in humour appreciation with repeated exposures to puns. The second study also investigated if humour on repetition depends on how humour is measured (as the emotional "mirth" experience of humour or according to a "cognitive appraisal" of the quality of the item; Gavanski, 1986). The challenge of comprehending incongruities (how long it takes

participants to "get" the humour in each pun), and the duration of time it took participants to provide humour ratings were also examined as predictors of changes in humour ratings with repetition.

In study 3, participants were asked to elaborate on the implied concepts in written puns. They were asked to list as many words as possible that come to mind when considering the two concepts from each pun. The number of elaborations that could be readily provided by participants and the duration of time they were willing to spend on this task for each pun were studied as predictors of humour appreciation. Latent semantic incongruity, familiarity with a pun from a prior occasion and the presence of aggressive content were investigated as predictors of humour from elaboration. Participant attention to semantic dissimilarity or semantic similarity was experimentally manipulated as independent variables of interest.

Chapter 2

2 Study 1: Latent Semantic Incongruity and Humour Appreciation

Semantic incongruity was operationalized here as the semantic dissimilarity between the alternate possible meanings implied by the key polysemous word in each written pun using Latent Semantic Analysis (LSA; Landauer, Foltz & Laham, 1998). Latent Semantic Analysis is analogous to the domain interaction approach, in that it can provide estimates of semantic relatedness between contrasted words or between text passages. Unlike the domains-interaction approach, LSA is purely mathematical, requiring no additional participant ratings to provide estimates of semantic similarity. Latent Semantic Analysis assumes that words with similar semantic content are likely to co-occur in the same passages of text within a sufficiently large digital corpus (a representative sample of text including books, articles, written content, or text passages). The currently reported studies used the default corpus offered by LSA which contains a collection of general readings up to the 1st year of college. Prior research suggests that LSA can provide estimate of semantic similarity in a way that is analogous to human judgments of semantic similarity in the context of studies of vocabulary recognition (Landauer & Dumais, 1997), and for semantic priming effects (Günther, Dudschig, & Kaup, 2016).

Landauer, Foltz and Laham (1998; http://lsa.colorado.edu) explain that LSA creates a matrix of co-occurrences for words (each row in the matrix represents a word) and text passages (passages are represented in each column) in a large digital corpus. Latent Semantic Analysis performs singular value decomposition (a form of factor analysis) over the matrix such that each word and passage is represented as a vector in high-dimensional space where each dimension can be considered a latent semantic factor. Latent Semantic Analysis also performs a pre-processing step in which the overall distribution of a word over its usage context is taken into account. This pre-processing step weights estimates for polysemous words (those with multiple meanings, such as in homographs) towards the meaning as it is most frequently used in the corpus. The similarity between words (each word being described as being an average of the meaning of all passages in which it appears), or between text passages (each passage being a kind

of average of the meaning of all words in the passage of interest) is computed as a cosine estimate of the similarity between the word or text passage vectors in high dimensional space. The cosine estimate can range from 0 to 1 with higher values indicating greater similarity. For example, the words "doctor" and "medicine" have a cosine similarity estimate of .74 given the relatively high likelihood that the two words appear in the same semantic context of the referenced corpus, whereas the words "doctor" and "flag" have a cosine similarity estimate of .01 suggesting that these two words are unlikely to co-occur in the same semantic context. It is important to note that in the current series of studies, estimates of the latent semantic incongruity between passages were reverse coded from LSA estimates (1.0 – LSA score), such that low semantic similarity estimates from LSA represents high levels of semantic incongruity (semantic dissimilarity). It is predicted here that latent semantic incongruity should be significantly positively correlated with participant humour ratings.

2.1   Polysemy Problem

It is potentially problematic to get an accurate assessment of semantic incongruity between the alternative meanings of polysemous words using co-occurrence models such as LSA. Latent Semantic Analysis examines the co-occurrence of words based on their overall usage, without taking into account syntactic or contextual constraints (although it is weighted to the dominant usage of the polysemous word in the corpus).

McHugh and Buchanan (2016) made an interesting methodological contribution to address the polysemy problem in their study of semantic ambiguity processing in homograph based puns. They used Durda and Buchanan's (2008) WINDSORS co-occurrence approach to measure latent semantic similarity in puns.[2] The primary distinction of their approach is that it can better account for high frequency words (words that are more frequently used in a given corpus would have inflated semantic similarity estimates). McHugh and Buchanan addressed the polysemy problem by disambiguating

---

[2] The WINDSORS approach to estimating semantic co-occurrence was not publicly accessible at the time studies 1-3 were conducted.

alternative meanings in their homograph based pun items by comparing two synonyms appropriate to the alternate implied senses of the homograph in each pun. For example, in the pun "those who play team sports usually have a ball", "ball" was identified as the homograph and "bat" and "fun" were prepared as synonyms appropriate to the two implied meanings. They compared the semantic similarity of the homograph to both of the synonyms using the WINDSORS approach. In their experiment, they presented entire puns to participants followed by one of the two synonyms as an experimental probe in a lexical decision task (participants were asked to determine if the presented word was a valid English word). When primed by a related pun, both synonyms had faster responses in comparison to responses for unrelated words (although the effect was larger for the synonym that was more strongly semantically related to the homograph). This indicates that both meanings of the homograph were activated and at play in comprehending the pun. McHugh and Buchanan's (2016) primary goal was to study semantic ambiguity processing and to provide converging validity for the WINDSORS approach. They did not ask participants to provide humour ratings for their pun items. In contrast, the primary goal of the current study was to examine semantic incongruity as a factor in humour processing and appreciation.

2.2   Dictionary Definition Approach

The present study examined semantic relatedness as an operationalization of incongruity in puns using a novel approach to address the polysemy problem. Latent Semantic Analysis passage vectors were created using dictionary definition entries that describe the alternative word meanings implied in written pun items. This is based on the assumption that definitions appropriate to the alternate implied word meanings of polysemous words are also necessarily text passages containing words that are highly semantically related to the appropriate implied sense of the concept. That is, a dictionary definition passage that describes the appropriate sense of a polysemous word should contain words that are each highly semantically associated with that implied concept. Given that LSA passage vectors provide similarity estimates judged over all of the words within each passage, dictionary definitions as passage vectors should provide robust estimates for the appropriate implied concept.

The dictionary definition approach was developed because single word synonyms chosen to represent alternate implied meanings can still also be polysemous, and because there was no clear and consistent approach to select a single word that is ideally semantically related and distinct to the appropriate senses of the polysemous word in puns. Selecting dictionary definitions that describe alternate implied senses as text passage vectors provided a consistent approach for vector selection. For example, for the pun "a cardboard belt would be a waist of paper", the dictionary definitions for "waist: the middle part of your body" and "waste: an action or use that results in the unnecessary loss of something valuable wasting a resource" were compared in LSA producing an LSA cosine estimate of .49 and therefore a latent semantic incongruity of estimate of .51 (1-LSA). Dictionary definition entries appropriate to the two concepts in each pun (according to the sense as used in each item) were taken from either the Miriam Webster online dictionary or from dictionary.com based on which volume appears to provide a more appropriate or representative entry for the concept as used in each pun (see Appendix B for further examples). Definitions were selected based on experimenter judgment and based on feedback from my supervisor and committee.

A major additional advantage of the present dictionary definition LSA co-occurrence approach, as compared to the domains interaction approach used by Hillson and Martin (1994) is that it allowed for the study of real-world examples of humour, namely written puns. Humour ratings for puns should be, on average, higher than those reported when using artificial metaphor stimuli. There should also be a greater range of humour ratings: on average some puns should prove to be very funny, while others may prove to be very unfunny. Puns as stimuli should also allow for greater variability in concepts as opposed to the artificial metaphor stimuli used in the domains-interaction approach. Hillson and Martin (1994) used 250 artificial-metaphor type jokes (created from 26 nouns and 6 domain-names; of the type "A is the B of A's domain"). In contrast, 300 ecologically valid written puns were used as stimuli for the current study and latent semantic incongruity estimates were prepared for each pun item.

2.3 Hypotheses

2.3.1 Latent Semantic Incongruity

Each of the aforementioned explanatory models of humour afford a central importance to incongruity for humour appreciation. Therefore, latent semantic incongruity estimates should be positively associated with participant ratings of funniness for pun items.

2.3.2 Familiarity

The incongruity-resolution (Suls, 1972) and comprehension-elaboration models (Wyer & Collins, 1992) both predict that a moderate level of comprehension difficulty should be important to humour from semantic incongruity in written puns. Familiar puns should not present a meaningful challenge for comprehension. Therefore, familiarity with a pun from a prior occasion should be negatively associated with humour ratings. Further, a low level of familiarity with a pun from a prior occasion should be important to humour produced from latent semantic incongruity. Puns with high semantic incongruity should be more humorous than puns with low semantic incongruity only for puns with a low level of familiarity from a prior occasion.

2.3.3 Aggressive Content

Misattribution theory predicts that aggressive humour is more enjoyable when it contains incongruous content (allowing people to believe they are enjoying the incongruities and not the aggression; Zillmann & Bryant, 1980). The incongruity-resolution based salience hypothesis holds that aggressive content should facilitate resolution and therefore increase humour produced from incongruous content (Suls, 1977; Goldstein, Suls & Anthony, 1972). These two theories both predict that aggressive content should be positively associated with humour, and that latent semantic incongruity and aggression should interact positively to further enhance participant humour ratings.

2.3.4 Pun Identification

The fluency account (Topolinski, 2014) predicts that humour appreciation is produced by rapid and easy comprehension of incongruities. Therefore, greater accuracy and shorter durations of time necessary for participant identification of items as written puns (as

opposed to non-pun control items) should be positively associated with humour. In contrast, the incongruity-resolution (Suls, 1972) and comprehension-elaboration (Wyer & Collins, 1992) models both hold that a moderate duration of processing should be positively associated with humour. A non-linear, inverted-u shaped, relationship between funniness ratings and pun identification duration would be supportive of this alternative hypothesis.

2.4 Method

2.4.1   Participants

One hundred and ninety-eight participants were recruited from the undergraduate pool at the University of Western Ontario. Of this total number, 18 participants were removed from analyses as outliers due to low item-total correlations with average scores across participants (less than .2 was used as a cut-off according to reliability analyses). Of the 180 participants included in the reported analyses, 50 were male and 130 were females (mean participant age = 18.72, SD = 3.14). Participants either spoke English as their first language (163 participants) or have been speaking English as their primary language for at least 10 years (17 participants; mean years of English experience = 12.71; SD = 3.49).

2.4.2   Materials and Procedure

Study 1 participants were asked to rate one of four lists of 100 items (of 180 participants, 44 participants were assigned to complete list 1, 46 completed list 2, 45 completed list 3, and 45 completed list 4). There were 75 puns and 25 non-pun controls in each list; thus, in total, across the four lists there were 300 pun items and 100 control items. The control items were created by substituting the polysemous word in actual puns with an unambiguous synonym that would be consistent with only one possible interpretation of the sentence (for a full list of pun and control items used in study 1, see Appendix A).

For each presented item, participants were first asked to complete the pun identification task, which consisted of identifying whether each item was either a pun, or if it was not a pun (in the case of control items). Following the pun identification task, participants were asked to rate each item on 7-point Likert type scales for: funniness, cleverness, the clarity

with which they understood the item, familiarity with an item from a prior occasion, the extent to which each item is a categorically good example of a pun, and for the presence of aggressive content.

The puns used in this study were collected from public submission pun databases available at punsandjokes.com and punoftheday.com. A corpus of 2000 puns was collected from these online resources: 100 homophone, 100 homograph and 100 rhyme-based puns were selected from this larger database. Puns with racist, sexist, offensive, or excessively aggressive content were not included in the current study to minimize ethical concerns. Puns from each of the three pun-type categories (homophone, homograph, rhyme) were selected with the primary goal of obtaining, as much as possible, variability in funniness ratings (items were selected which appeared to be low, moderate, and high in funniness according to the intuition of the primary researcher, but with feedback and input from Dr. Rod Martin, Dr. Albert Katz, and from a research assistant (Lisa King).

Each list of items contained a different set of: 25 puns based on homographs, 25 puns based on homophones, 25 puns based on rhymes, and 25 control items. Control items were constructed such that the semantically ambiguous key word of the pun was replaced with a non-ambiguous synonym that is consistent with the topic meaning of the sentence (34 control items were based on homophone puns; 33 on homographs and 33 were based on rhyme). For example, a control item was created by substituting the pun word in "As a matter of *flat*, he lives on the 2nd floor" with an unambiguous synonym as in "As a matter of *fact*, he lives on the 2nd floor". Four lists of 100 items were structured such that participants would not see both the control and actual pun version in the same list.

The study used a custom internet survey platform developed by Dr. Rod Martin. Participants were able to complete the study at a time and place of their choosing. Participants were first asked to provide general demographic details: self-reported age, gender, whether English was their first language (yes/no) and if English was not their first language, they were asked to indicate the number of years for which they have been speaking English.

Participants were assigned to rate one of four lists of 100 items. The items in each list were presented one at a time for ratings and in a random order for each participant. Participants were first asked to judge whether each item is a pun or is not a pun. The duration of time necessary to judge if an item is a pun (in milliseconds) and accuracy of the pun identification judgment was measured (dichotomously coded as being 0 when erroneous and 1 when correct). Next, each of the 100 items on each list was rated according to six 7-point Likert-type scales [each ranging from not at all (0) to extremely (6)]. Participants were allowed to select "no answer" if they did not feel comfortable providing pun ratings for a given item. Participants were asked to rate the extent to which each item is: "a good example of a pun", "aggressive", "funny", "clever", "familiar from a prior occasion", and "clearly understood". Participants were assigned to rate these six dependent variables scales in a randomized order for each run of the experiment.

2.4.3   Latent Semantic Incongruity

For all Latent Semantic analysis estimates reported in this dissertation, the default LSA settings were used for pairwise comparisons of concepts from written pun items: the analyses used the corpus of general readings up to 1st year of college; calculated to 300 factors). Dictionary definitions that were used as passage vectors for the LSA analysis were taken from either Miriam Webster online dictionary (www.merriam-webster.com) or dictionary.com. Dictionary definitions for the alternate meaning were prepared for the two implied concepts from the key ambiguous word in each item. For example, "Pitch: to hurl or throw (something); cast" vs. "Pitch: to aim to sell (a product) to a specified market or on a specified basis") in the pun "A baseball player can sell himself to a new team if he has a good pitch". Dictionary definitions of alternate meanings in puns were prepared for each of the 300 puns used in the current study (100 homophone, 100 homograph, and 100 rhyme based puns). The dictionary definition approach was applied to homophone and rhyme puns for the sake of consistency and to permit comparison of latent semantic incongruity effects using the same method between pun-types (latent semantic incongruity could potentially be calculated for rhyme and homophone based on the semantic dissimilarity between the single polysemous word and single alternate implied word; also, single word comparisons between concepts can still be polysemous).

Control items could not be included in the LSA analysis, as they no longer contain clear polysemous pun-based wordplay.

2.4.4   Data Preparation

2.4.4.1   Missing Data

Participant responses for the pun identification task that took longer than 25000 milliseconds (approximately two standard deviations above the mean) were discarded and replaced with the mean across other participants for that item (7.4% of the 18000 pun identification durations were replaced). Missing data due to participant selection of the "no answer" option was also replaced with the mean for that scale across participants. Missing data replacement with the mean across participants due to selection of "no answer" was relatively rare for ratings of funniness (764 of 18000 cases; 4.24%), cleverness (805 of 18000 cases; 4.47%), clarity of understanding (675 of 18000 cases; 3.75%), familiarity (840 of 18000 cases; 4.67%), aggression (1121 of 18000 cases; 7.40%), and the extent to which an item is a good example of a pun (807 of 18000 cases; 4.48%).

All inferential statistical analyses were conducted on data that was averaged across items. Participant responses were averaged across items for all likert-scale ratings: for funniness, cleverness, clarity of understanding, familiarity,[3] aggressiveness and the extent to which each item is a categorically good example of a pun.

2.4.4.2   Internal Reliability

Reliability analyses were conducted with participant response data transposed such that participants were treated as items on a scale. Thus, item-total correlations could be examined as an estimate of the extent to which each participant's responses are consistent with the average participant response for each list. Data from participants with low item-

---

[3] Average familiarity over items should be viewed as the likelihood that a pun will be known from a prior occasion or as the average frequency at which an item is used in the current cultural context

total correlations (lower than .2) according to any one scale were removed from analysis as atypical outliers. Using this cut-off criterion, responses from 18 participants were removed from further analyses.

Cronbach's alpha estimates were prepared to assess participant reliability (with 180 of 198 participants included in final analyses) for all measured variables and across each of the four item lists. For the complete set of Cronbach's alpha reliability estimates, refer to Table 1. With the exception of aggression ratings for list 3 (which was still satisfactory at .69), reliability estimates were acceptably high. For all other dependent variables over the four item lists, Cronbach's alpha estimates ranged from .80 to .95, indicating a high level of internal consistency between participants for ratings of the written puns and control items.

2.5   Results

2.5.1   Factor Analysis

Prior to conducting statistical tests to evaluate study 1 hypotheses, a principal component factor analysis was conducted on pun items (control items were not included). The mean scores for all dependent variables (except for hypothesized moderators; aggression and familiarity) were included in the factor analysis: funniness, cleverness, clarity of understanding, and the extent to which the item is a good example of a pun. The mean and standard deviation for each of these variables broken down by item type (homophone, homograph, rhyme based puns or control items) is presented in Table 2. Using a criterion of eigenvalues greater than one, the factor analysis suggested a single latent factor solution could best describe the variables, consisting of high loadings on each of the four included variables. The single factor was able to account for 91.78% of variability in the four scales and was labeled the as being an overall "effective humour" factor. The extent to which each pun item loads on the effective humour factor was saved using the regression method, to be used as a primary outcome variable of interest for subsequent analyses.

Table 1: Cronbach's alpha reliability estimates for participants on each study 1 variable broken down by item list.

	List 1 α	List 2 α	List 3 α	List 4 α
Funniness	0.89	0.91	0.90	0.97
Cleverness	0.89	0.97	0.91	0.88
Good Example of a Pun	0.89	0.88	0.91	0.89
Clarity of Understanding	0.95	0.91	0.91	0.92
Familiarity	0.87	0.92	0.85	0.80
Aggression	0.87	0.92	0.66	0.92
Identification Accuracy	0.85	0.89	0.89	0.88
Identification Duration	0.84	0.85	0.81	0.90

Table 2: Mean over items and standard deviation for study 1 variables.

Scale	Homophone Mean (*SD*)	Homograph Mean (*SD*)	Rhyme Mean (*SD*)	Control Mean (*SD*)
Clarity	4.68 (0.82)	4.8 (0.58)	4.29 (0.77)	2.24 (0.64)
Clever	3.33 (0.78)	3.21 (0.68)	2.87 (0.66)	0.57 (0.44)
Funny	3.04 (0.78)	2.92 (0.65)	2.58 (0.68)	0.72 (0.49)
Good example of a pun	3.34 (0.81)	3.2 (0.69)	2.83 (0.67)	0.62 (0.47)
Familiarity	2.69 (0.69)	2.71 (0.49)	2.35 (0.57)	1.13 (0.42)
Aggressive	1.3 (0.61)	1.24 (0.59)	1.17 (0.46)	0.52 (0.4)
Identification Accuracy	0.84 (0.15)	0.84 (0.13)	0.78 (0.15)	0.85 (0.12)
Identification Duration	6752.08 (1496.79)	6119.69 (1181.68)	7276.06 (1563.72)	8172.32 (1660.58)
Latent Semantic Incongruity	.49 (.14)	.49 (.15)	.49 (.14)	

Note: Averages were reported broken down by item type (homophone, homograph, rhyme based puns or control items). Outcome variables were on scales from 0 (not at all) to 6 (extremely). Pun identification accuracy was dichotomously coded as 0 (incorrect identification) or 1(correct identification). Pun identification duration is reported in milliseconds.

2.5.2   Comparison of Different Types of Pun Mechanisms

Orthographic (entries available for N=281 of 300 puns) and phonological frequency (entries available for N=271 of 300 puns) estimates were prepared according to the key ambiguous word in each pun as spelled, or according the dominant or first implied meaning of the word from the pun sentence using wordmine2 (Durda & Buchanan, 2006). Across all pun items there was no significant correlation between effective humour factor loadings and orthographic frequency $r(279) = -.05$, *ns* or phonological frequency estimates $r(269) = -.07$, *ns*. The complete list of frequency estimates are presented in Appendix A.

Exploratory analyses using independent samples *t*-tests were conducted to investigate whether there were significant differences between study 1 items based on the wordplay mechanism used to create humour (homograph, homophone or rhyme based puns or control items). Effective humour factor loadings, aggression, familiarity, identification accuracy and identification duration were investigated as variables of interest.

The only significant difference between homophone (based on the phonological similarity of the key polysemous word) and homograph based puns (which make use of both phonological and orthographical overlap) was that homograph based puns were identified significantly faster than homophone based puns $t(198) = 3.32, p < .001$. Both homophone and homograph based puns had significantly higher effective humour factor loadings than rhyme based puns [homophone $t(198) = 4.50, p < .001$; homograph $t(198) = 4.22, p < .001$]. Homophone and homograph based puns were identified more accurately than rhyme based puns [homophone $t(198) = 2.90, p < .01$; homograph $t(198) = 2.70, p < .01$]. Homophone and homograph based puns were also identified significantly more rapidly than rhyme based puns [homophone $t(198) = 2.42, p < .05$; homograph $t(198) = 5.90, p < .001$]. Homophone and homograph based puns were both more likely to be familiar from a prior occasion [homophone $t(198) = 3.80, p < .001$; homograph $t(198) = 4.77, p < .001$] than rhyme based puns.

2.5.3   Latent Semantic Incongruity

The latent semantic incongruity between the two implied meanings in each of the puns

Table 3: Bivariate correlation between study 1 variables.

Scale	1	2	3	4	5	6	7	8	9	10
1. Incongruity	-	.29**	.32**	.31**	.18**	.29**	.19**	.21**	.21**	0.01
2. Effective Humour Factor	-	-	.97**	.98**	.90**	.98**	.42**	.84**	.86**	-.40**
3. Funniness	-	-	-	.96**	.81**	.95**	.40**	.79**	.80**	-.36**
4. Cleverness	-	-	.56**	-	.81**	.97**	.43**	.78**	.82**	-.33**
5. Clarity	-	-	.28**	.55**	-	.84**	.34**	.86**	.85**	-.46**
6. Good Example of Pun	-	-	.34**	.84**	.33**	-	.43**	.82**	.85**	-.38**
7. Aggressive	-	-	.26**	.55**	0.15	.54**	-	.36**	.40**	-0.08
8. Familiarity	-	-	0.17	.60**	.83**	.47**	.21*	-	.71**	-.46**
9. Identification Accuracy	-	-	-.59**	-.91**	-.51**	-.77**	-.51**	-.49**	-	-.35**
10. Identification Duration	-	-	0.07	-0.01	-.32**	0.04	0.18	-.23*	-0.01	-

Note: Intercorrelations amongst pun items are presented above the diagonal, correlations amongst control items are presented below the diagonal. Correlations marked by (*) are statistically significant at $p < .05$, correlations marked by (**) are statistically significant at $p < .01$.

was significantly correlated with effective humour factor loadings $r(298) = .29$, $p < .001$, aggressiveness $r(298) = .19$, $p < .001$, familiarity with a pun from a prior occasion $r(298) = .21$, $p < .001$, and identification accuracy $r(298) = .21$, $p < .001$. The complete series of bivariate correlations between latent semantic incongruity, and effective humour factor loadings is presented in Table 3 (correlation amongst pun items presented above the diagonal, correlation amongst non-pun control items is presented below the diagonal).

2.5.4   Moderator Variable Analyses

Separate multiple regression moderation analyses were conducted to determine whether each of the hypothesized variables significantly moderates the association between latent semantic incongruity and the effective humour factor. Moderation analyses were conducted on average participant ratings for the 300 pun items (control items were not included). Each analysis used a stepwise approach: in the first step the main effect of latent semantic incongruity on effective humour factor loadings was established. In a second step, the moderator variable was added. In the final step of each analysis, the cross product of the moderator and latent semantic incongruity was added to assess for a significant moderating interaction effect. If a significant moderating interaction was identified, a post hoc simple test of slopes was conducted to study the interaction further (this post hoc test examines whether the slope from low to high values of an independent variable is significantly greater than zero; see Dawson & Richter, 2006).[4]

2.5.4.1   Aggression

At the first step, there was a significant effect of latent semantic incongruity on effective humour factor loadings $\beta = 2.03$, $p < .001$; $R^2 = .08$, $F(1, 298) = 27.26$, $p < .001$. At the second step, there was a significant main effect of both latent semantic incongruity $\beta = 1.53$, $p < .001$ and aggression $\beta = .67$, $p < .001$; $R^2 = .22$, F change $(1, 297) = 52.62$, $p < .001$. At the final step there was a significant main effect of both latent semantic

---

[4] Post hoc analysis of slopes was conducted using templates available at http://www.jeremydawson.co.uk/slopes.htm).

incongruity $\beta = 3.25$, $p < .001$ and aggression $\beta = 1.39$, $p < .001$ and there was a significant moderating interaction of aggression on the association between latent semantic incongruity and effective humour, $\beta = -1.37$, $p < .05$; $R^2 = .23$, F change $(1, 296) = 4.56$, $p < .05$.

To further examine the interaction, low latent semantic incongruity was defined as .20 (incongruity estimates could range from 0 to 1.0; based on approximately two standard deviations below the mean; refer to Table 2), high latent semantic incongruity was defined as .80 (based on approximately two standard deviations above the mean). Low aggression was defined as being 1, moderate aggression was defined as being 3 on a scale from 0-6; based on approximately 2 SD above the mean). There was a significant effect of semantic incongruity, but only for puns with a low level of aggression $t(299) = 4.70$, $p < .001$ (see Figure 1). There was also a significant effect of aggression, but only for puns with low latent semantic incongruity $t(299) = 4.91$ $p < .001$ (see Figure 2). That is, on average, low aggression was important for humour associated with semantic incongruity and low semantic incongruity was important for humour associated with aggression.

2.5.4.2   Familiarity

At the first step, there was a significant effect of latent semantic incongruity on effective humour factor loadings $\beta = 2.03$, $p < .001$; $R^2 = .08$, $F(1, 298) = 27.26$, $p < .001$. At the second step, there was a significant main effect of both latent semantic incongruity $\beta = .82$, $p < .001$ and familiarity $\beta = 1.34$, $p < .001$; $R^2 = .73$, F change $(1, 297) = 695.74$, $p < .001$. At the final step, there was a significant main effect of both latent semantic incongruity $\beta = 3.10$, $p < .01$ and familiarity $\beta = 1.76$, $p < .001$, and there was a significant moderating interaction of familiarity on the association between latent semantic incongruity and the effective humour factor $\beta = -.86$, $p < .05$; $R^2 = .73$, F change $(1, 296) = 5.33$, $p < .05$.

Once again, low latent semantic incongruity was defined as .20, high latent semantic incongruity was defined as .80. Low familiarity was defined as being 1 while high familiarity was defined as being 4 (on a scale from 0-6; values chosen based on approximately 2 SD above and below familiarity mean; refer to Table 2). There was a
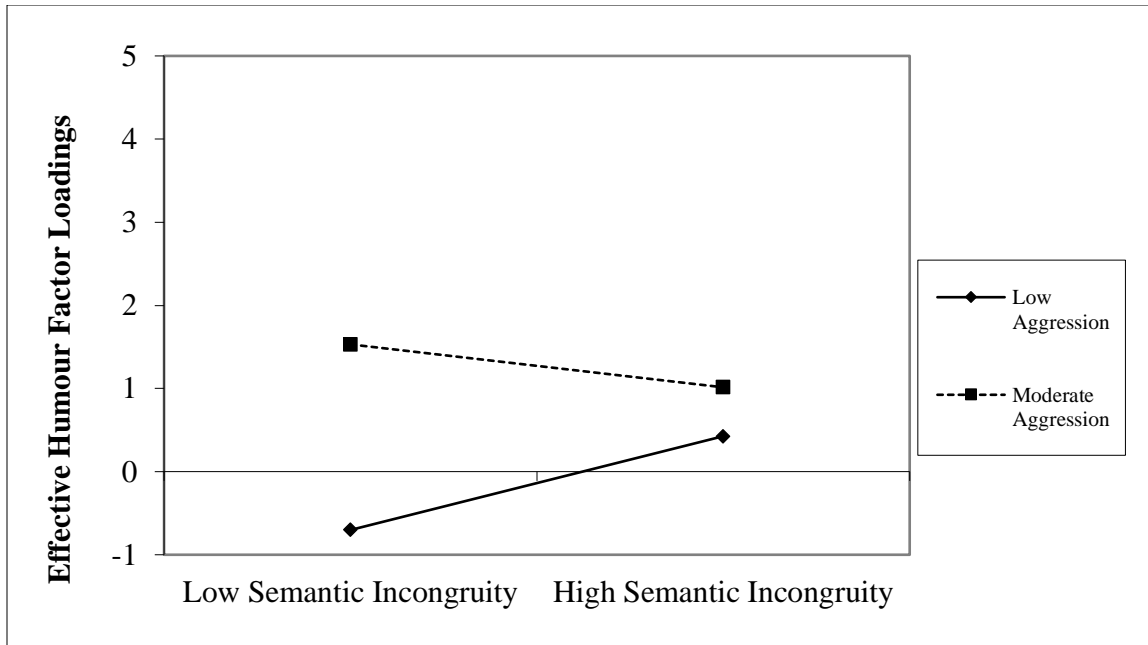
Figure 1: Moderating effect of aggression on the relationship between latent semantic incongruity and item loadings on the effective humour factor in study 1.

Puns with high latent semantic incongruity (.8) had higher loadings on the effective humour factor than puns with low latent semantic incongruity (.2) only for items with low aggression (set as 1; on a scale from 0 to 6). There was no significant effect of latent semantic incongruity for puns with a moderate level of aggression (set as 3; approximately 2 *SD* above and below mean ratings of aggressive content).
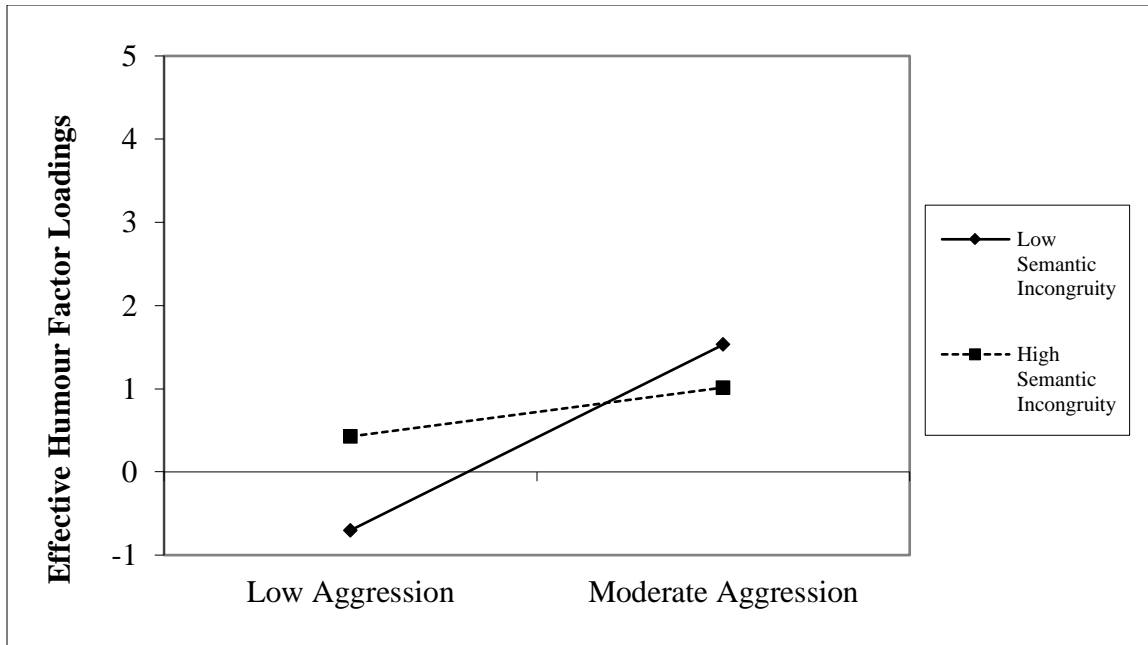
Figure 2: Moderating effect of latent semantic incongruity on the relationship between aggression and item loadings on the effective humour factor in study 1.

Puns with moderate aggression (set again as 3) had higher effective humour factor loadings than puns with low aggression (set again as 1) only for items with low latent semantic incongruity (set again as .2; high latent semantic incongruity set as .8).

significant effect of latent semantic incongruity only for puns with low familiarity $t(299)$ = 2.22, $p < .05$, but there was no effect of semantic incongruity for puns with a high level of familiarity. Puns with high familiarity were more humorous than puns with low familiarity regardless of the level of semantic incongruity (see Figure 3).

2.5.4.3   Exploratory Analysis: Familiarity and Aggression on Humour

An exploratory moderation analysis was conducted to investigate familiarity as a moderator of the association between aggression and the effective humour factor loadings. At the first step, there was a significant effect of aggression on the effective humour factor $\beta = .75$, $p < .001$; $R^2 = .18$, $F(1, 298) = 63.54$, $p < .001$. At the second step, there was a significant main effect of both aggression $\beta = .24$, $p < .001$ and familiarity $\beta = 1.30$, $p < .001$; $R^2 = .73$, F change $(1, 297) = 606.84$, $p < .001$. At the final step, there was a significant main effect of both aggression $\beta = .99$, $p < .001$ and familiarity $\beta = 1.63$, $p < .001$, and there was a significant moderating interaction of familiarity on the association between aggression and the effective humour factor $\beta = -.27$, $p < .001$; $R^2 = .74$, F change $(1, 296) = 11.18$, $p < .001$.

As in prior sections, low familiarity was defined again as being 1 while high familiarity was defined as being 4 (on a scale from 0-6). Low aggression was defined again as being 1, moderate aggression was defined as being 3. There was a significant effect of aggression only for items with low familiarity $t(299) = 4.62$, $p < .001$ and there was no significant effect of aggression for puns with a high level of familiarity. Puns with high familiarity were significantly more humorous than puns with low familiarity regardless of whether the item had aggressive content (see Figure 4).

2.5.4.4   Identification (Accuracy and Duration)

At the first step, there was a significant effect of latent semantic incongruity on effective humour factor loadings $\beta = 2.03$, $p < .001$; $R^2 = .08$, $F(1, 298) = 27.26$, $p < .001$. At the second step, the pun identification task variables were added as predictors (both identification accuracy and identification duration). There was as a significant main effect of latent semantic incongruity $\beta = .86$, $p < .001$, identification accuracy $\beta = 5.48$, $p < .001$
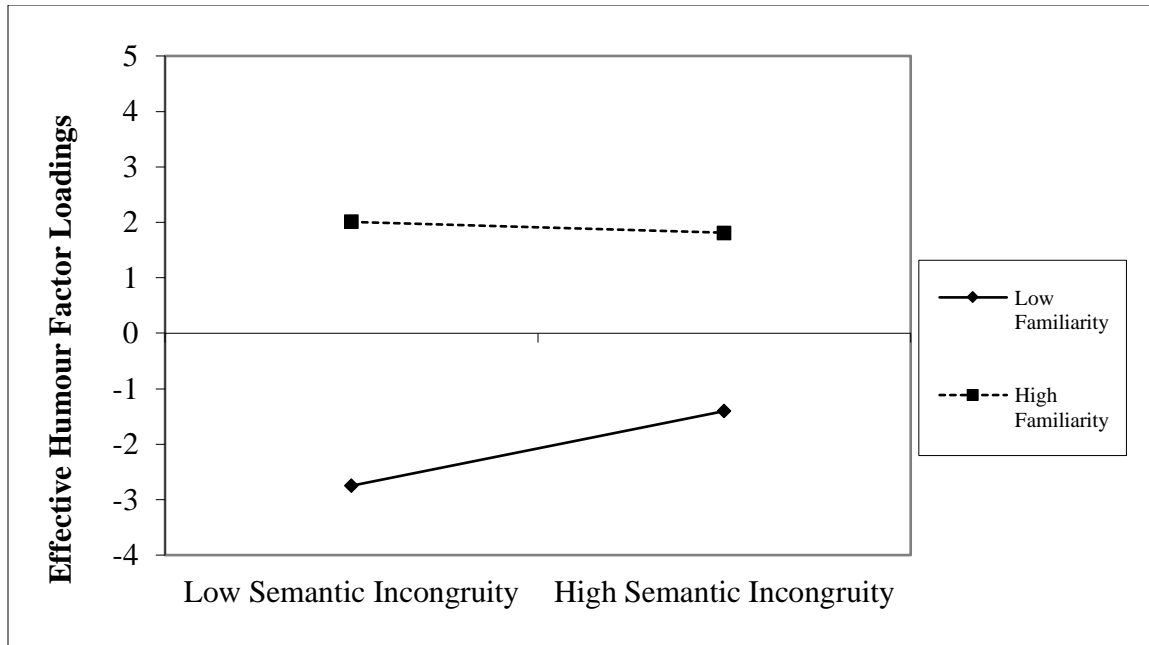
Figure 3: Moderating interaction of familiarity on the relationship between latent semantic incongruity and item loadings on the effective humour factor in study 1.

Puns with high latent semantic incongruity (set as .8) had significantly higher effective humour than to puns with low latent semantic incongruity (set as .2) only for items with a low level of familiarity (set as 1; high familiarity set as 4 approximately 2 *SD* above and below familiarity mean on a scale from 0 to 6).
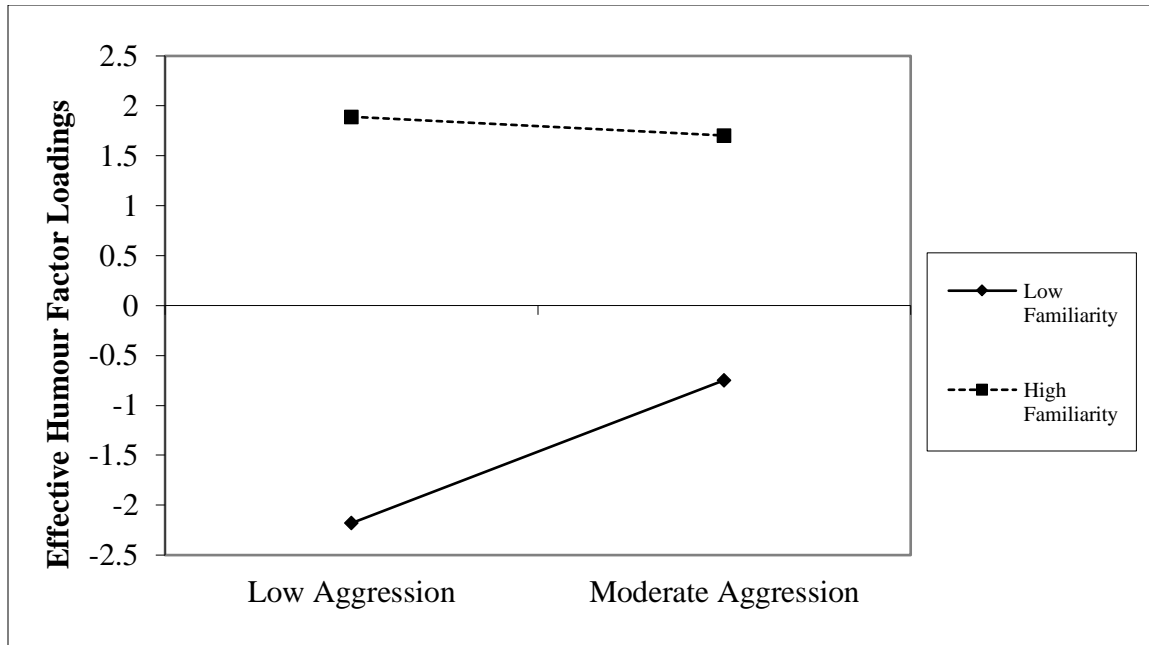
Figure 4: Moderating effect of familiarity on the relationship between aggression and item loadings on the effective humour factor in study 1.

Puns with moderate aggression (set as 3) had higher loadings on the effective humour factor than puns with low aggression (set as 1) only for items with low familiarity (set as 1; both variables on a scale from 0 to 6). There was no effect of aggression for puns with a high level of familiarity (set as 4).

and identification duration β = -0.000083, $p < .001$; R^2= .77, F change (1, 296) = 439.61, $p < .001$. At the final step, there was a significant main effect of identification accuracy β = 4.30, $p < .001$ and identification duration β = -0.00020, $p < .01$, but there were no significant main effect of latent semantic incongruity β = -2.79, ns. There was also no significant moderating interaction effect of identification accuracy β = 2.52, ns and no significant interaction for identification duration β = -2.79, ns on the association between semantic incongruity and effective humour. At the third step adding these terms also did not significantly improve model fit R^2= .77, F change (1, 294) = 1.79, ns. Therefore, the association between latent semantic incongruity and effective humour was not significantly influenced by identification accuracy or identification duration.

There was no statistically significant non-linear (quadratic; inverted-u shaped) association between identification accuracy, identification duration and the effective humour factor.

2.6 Discussion

The primary purpose of study 1 was to develop and test a novel measure of latent semantic incongruity in written puns. The reported analyses reveal several novel insights into the function of semantic incongruity in humour (operationalized as semantic dissimilarity using LSA), the linguistic mechanisms used to create puns, and provides novel insights into how people identify and evaluate humour in puns.

There was a high degree of consistency in participant ratings as assessed using Cronbach's Alpha. Although participants would likely have been bringing a wide range of personal preferences, experiences, and prior knowledge into the experiment, there was a high level of consistency in their humour ratings. Although it was not the focus of the studies reported in this dissertation, future research investigating individual differences in humour appreciation would benefit from considering participant preferences for different styles of humour (for a discussion see Martin, Puhlik-Doris, Larsen, Gray, & Weir, 2003), individual differences in phonological decoding ability (Jared & Bainbridge, 2017), or examining a broader sample of participants. For example, there may be greater variability between individuals in studies of humour appreciation with children (as in Zigler, Levine

& Gould, 1967; McGhee, 1976) or with adults with less language proficiency (for example as with novice second language learners).

The dependent measures in the current study were highly intercorrelated (funniness, cleverness, clarity of understanding and the extent to which each item is a categorically good example of a pun). Therefore, for the purpose of data reduction, these variables were combined using a factor analysis and assessed as a single "effective humour" latent factor. Item loadings on the effective humour factor were utilized as the primary outcome variable of interest to test predictions from empirical models of humour appreciation.

2.6.1   Latent Semantic Incongruity

Incongruity is frequently discussed as a variable of central importance to humour appreciation, but the concept has been challenging to consistently operationalize and measure for empirical study (Ritchie, 1999; 2004; Forabosco, 1992). Building on prior work by Hillson and Martin (1994), the current study developed a new approach to operationalize incongruity as semantic dissimilarity using latent semantic analysis (LSA; Landauer, Foltz & Laham, 1998). The novel "latent semantic incongruity" approach estimated the semantic incongruity between dictionary definitions appropriate to the two possible implied concepts in puns. This measure was positively associated with humour which provides support for a core hypothesis held by both the incongruity-resolution model (Suls, 1972) and comprehension-elaboration model (Wyer & Collins, 1992): that incongruity is important to humour appreciation. Moderator analyses further demonstrated that the relation between semantic incongruity and effective humour is dependent on having a low level of aggressive content and a low level of familiarity with a pun from a prior occasion. The implications of these moderating interactions are discussed in greater detail in the next section.

2.6.2   Pun Mechanisms

Exploratory analyses compared participant ratings for puns based on homophones, homographs and rhymes. Current results suggest that phonological overlap (but not necessarily overlap in spelling) is important to humour from puns. The only statistically significant difference between homophone and homograph based puns was that

homograph based puns were identified as puns significantly faster than homophone based puns. Overlap in spelling may not be important to humour from written puns given that all homograph based puns used in the current study were also homophones (with an overlap in *both* spelling and sound) and yet there were no differences in effective humour between homophone and homograph based puns. The finding that homophone and homograph based puns were both significantly more humorous than puns based on rhymes further supports this interpretation (given that rhymes are based only on a degree of phonological similarity).

Homophone and homograph based puns were more typical and readily recognizable as puns than rhyme based puns. They were identified as puns more accurately and rapidly than rhyme based puns. Further, homophones and homographs were also rated as being more familiar on average; which indicates that homophone and homograph based puns may be experienced more frequently in everyday life.

2.6.3   Moderator Analyses

2.6.3.1   Aggression

Aggressive content in puns was a strong positive associate of effective humour. The current study did not use puns with excessive levels of aggression, so current findings should be generalized only to puns with low to moderate levels of aggressive content,[5] however, there was also a significant moderating interaction such that the effect of semantic incongruity was only visible for puns with a low level of aggressive content and the effect of aggression was only visible for puns with a low level of semantic incongruity. Based on current findings, it can be speculated that semantic incongruity and aggression may be mutually incompatible as potential sources of humour. The finding that there is no effect of aggression for puns with high semantic incongruity challenges

---

[5] Sexist, racist, offensive, and excessively aggressive puns were not used in the current study. Therefore, results can only be generalized to stimuli with low to medium levels of aggressive content.

the core assumption from superiority theory that aggression is necessary for humour (Gruner, 1997). These findings are also problematic for models which argue that aggression and semantic incongruity in humorous stimuli should mutually facilitate the experience of humour. According to the salience hypothesis, aggressive content should have enhanced humour from semantic incongruity by facilitating resolution (Goldstein, Suls & Anthony, 1972; Suls, 1977) and according to misattribution theory, incongruity should have enhanced humour from aggressive content by making offensive content seem more socially permissible (Zillmann & Bryant, 1980). The hypothesized mutually facilitative effects were not obtained. Semantic incongruity did significantly interact with humour from aggressive content, but not as hypothesized by these models.

Overall, aggression in written puns appears to be a potential source of humour that is distinct from semantic incongruity. Prior research by Ruch et al. (Ruch, 1992; Ruch 1981, 1984, Ruch & Hehl, 1998) consistently found three distinct categories of humour: including humour from incongruity-resolution, humour from sexual themes and nonsensical humour (that cannot clearly be explained). Results from the current study suggest that aggressive content may yet be a meaningfully distinct category of humorous stimuli. In support of this argument, novelty (as operationalized by a low level of familiarity with a pun from a prior occasion) was important to humour from aggressive content: puns with a moderate level of aggressive content were more humorous than puns with a low level of aggressive content only for puns with a low level of familiarity from a prior occasion. As with humour from semantic incongruity, humour from novel aggressive content may need to be resolved (explained or understood). The positive association between aggressive content and humour only for low familiarity puns is also consistent with Herzog et al.'s hypothesis that surprise or shock (and therefore novelty) can account for why people enjoy aggressive, "sick", or offensive jokes (Herzog & Bush, 1994; Herzog & Karafa, 1998).

2.6.3.2   Familiarity

Items with high latent semantic incongruity were more humorous than items with low latent semantic incongruity only for puns with a low average level of familiarity from a prior occasion. This finding is consistent with predictions from the incongruity-resolution

model (Suls, 1972), which emphasizes the importance of surprise and novelty. However, there was also an unexpected overall strong positive correlation between humour and familiarity with a pun from a prior occasion. That is, regardless of the level of latent semantic incongruity or the level of aggression in an item, puns with a high level of familiarity from a prior occasion were always significantly more humorous than puns with a low level of familiarity. It is also important to note that the directionality of this finding is unclear: familiarity with a pun may enhance humour (as would be predicted by the fluency account; Topolinski, 2014) or the most effective puns may be more likely to be recognized or recalled from prior exposures (consistent with what would be predicted by Suls, 1972; 1975). Although novelty proved to be important to humour from both latent semantic incongruity and aggression, familiarity was actually an overall relatively more important associate of humour: there was always a significant strong positive association between humour and familiarity ratings regardless of the level of semantic incongruity or aggression. To further investigate the function of familiarity in humour appreciation, familiarity will be experimentally manipulated as a variable of interest in study 2.

The finding that participants rated familiar puns as being more humorous appears to be counter-intuitive at first glance. However, anecdotally, in "running gags" jokes can actually increase in humour with repetition and people can re-watch a favorite comedy sit-com series several times and still find it funny. Prior empirical studies have also demonstrated that jokes with more easily predictable punch lines are significantly more humorous (Kenny 1955; Pollio & Mers, 1974; Topolinski, 2014). While the incongruity-resolution model (Suls, 1972) predicts that a moderate level of comprehension difficulty should be associated with humour, these results suggest that the more rapid processing and easier challenges associated with familiar stimuli should make repeated stimuli more humorous.

2.6.3.3   Identification Accuracy and Duration

The incongruity-resolution (Suls, 1972) and comprehension-elaboration (Wyer & Collins, 1992) models predicted that that there should have been a non-linear (inverted-u shaped) relation between comprehension difficulty and humour ratings. Humour should increase

with the level of comprehension difficulty presented by a pun to an ideal point, and then humour should decrease if comprehension becomes too difficult. However, there were no significant nonlinear associations between identification accuracy or identification duration and humour appreciation. Identification accuracy and duration also did not significantly moderate the association between latent semantic incongruity and humour. That is, semantic incongruity was associated with humour regardless of processing difficulty or fluency.

Consistent with predictions from the fluency account (Topolinski, 2014), both identification accuracy and rapid identification duration were significantly positively correlated with effective humour factor loadings. Further, according to correlational analyses, puns with high latent semantic incongruity could be more accurately (but not more rapidly) identified as written puns. Results from the identification task are potentially limited by the fact that puns, as a category of humorous stimuli, are relatively easy to correctly identify and by the fact that full comprehension of a pun may not have been necessary to identify an item as a pun. That is, participants may have been able to accurately recognize an item as a pun without yet fully understanding the humour. These limitations will be addressed by including a task that can more directly assess comprehension difficulty in study 2.

2.6.4   Study 1 Summary

Consistent with the incongruity-resolution (Suls, 1972) and comprehension-elaboration models (Wyer & Collins, 1992) which both emphasize the importance of incongruity in humour, the new measure of latent semantic incongruity developed here was positively associated with participant humour ratings. The core prediction from the incongruity-resolution model (Suls, 1972) and comprehension-elaboration model (Wyer & Collins, 1992) that incongruity should be positively associated with humour only for puns with a low level of familiarity from a prior occasion was also supported. However, incongruity-resolution and comprehension-elaboration could not clearly account for the effects of aggression, familiarity, and processing difficulty (according to pun identification accuracy and duration). It was predicted that aggression and incongruity should interact productively for humour. However, aggression appeared to function as a potential source

of humour that was distinct from humour based on semantic incongruity. The apparent incompatibility of humour from semantic incongruity and aggressive content was also not clearly predicted by superiority theory (Gruner, 1997), misattribution theory (Zillmann & Bryant, 1980) or the salience account (Goldstein, Suls & Anthony, 1972). Although low familiarity (and therefore surprise and novelty) was important to humour from both semantic incongruity and from aggression, the overall strong positive association between familiarity and effective humour factor loadings was also not expected. The overall positive association between rapid and accurate pun identification and humour ratings is instead more supportive of the prediction from the fluency account (Topolinski, 2014), that quick and easy comprehension should be predictive of humour appreciation.

Chapter 3

3 Study 2: Tell Me One I Haven't Heard Before

Just as incongruity has been challenging to operationalize for experimental study, the "resolution" of incongruity has been similarly challenging to define. Prior studies have operationalized resolution as the explanatory mechanisms in humorous stimuli that can be used to comprehend incongruity. Explanatory mechanisms can include elements such as the punch line of a joke or the semantic similarities between contrasting concepts in humorous comparisons (as with humorous metaphors in Hillson & Martin, 1994). In accordance with this operationalization, Attardo (1994; see also Attardo & Raskin, 1991) argued that mere recognition of the connotative semantic association between two possible meanings in written puns should be sufficient as resolution. Resolution has also been defined as the cognitive process by which incongruities are understood (for example as a process of re-interpretation or due to violated expectations; Suls, 1972; Wyer & Collins, 1992; see also Ritchie, 2004, 2009). These two definitions both implicitly view resolution as the problem-solving process by which incongruities are explained. To further study resolution in humour appreciation from semantic incongruity, the current study operationalized resolution as the measurable changes in humour assessment over repeated exposures (after initial comprehension) to humorous written puns. Humour produced by semantic incongruity was predicted to decrease (partially or completely) with repeated exposures to a pun. If resolution of novel incongruities is important to humour appreciation, then only comprehension challenge (the difficulty of explaining incongruities) at first exposure (and not repeated exposures) should be predictive of humour. Changes in the function of semantic incongruity with repetition in humour appreciation is hypothesized to be indicative of changes in assessment due to resolution.

3.1 The Repetition Problem

The primary goal of study 2 was to investigate semantic incongruity as a predictor of changes in participant humour ratings for repeated puns. Prior studies have most frequently found that humour ratings tend to decrease with repeated exposures

(Goldstein, 1970b as cited in Suls, 1975; Pistole & Shor, 1979, Gelb & Zinkhan 1985; Gavanski, 1986). However, some jokes, even though they are clearly well remembered, can continue to produce humour even when experienced several times. In some circumstances, such as in a running gag, humour can actually increase with repeated exposures. Prior studies have also found that humour ratings can persist with repeated exposures (Suls, 1975; Zhang & Zinkhan, 1991). Although incongruity-resolution emphasizes the importance of surprise, higher levels of predictability of punchlines has been shown to be positively associated with humour (Kenny, 1955; Pollio & Mers, 1974; Topolinski, 2014). Results from study 1 further suggest that familiarity with puns from a non-specific prior occasion may actually increase humour appreciation. There was a significant effect of semantic incongruity only for puns with a low level of familiarity, but familiarity with a pun from a prior occasion was found to be an overall strong positive associate of participant humour ratings (for consistent evidence see also Schick, McGlynn & Woolam, 1972). Persistent or increased humour over repeated exposure to humorous stimuli poses a considerable challenge for the incongruity-resolution model, which emphasizes the importance of incongruity comprehension in humour appreciation (Suls, 1972). Repeated exposures to humorous stimuli should not present the same level of challenge for comprehension or surprise as experienced at first exposure; and yet repeated stimuli can still be humorous.

3.2   Explanations for Humour on Repeated Exposures

Several explanations have been proposed to address the repetition problem from an incongruity-resolution perspective: Suls (1972; 1975) hypothesized that the problem could be explained by mere-exposure effects, re-interpretation of familiar items based on discovery of previously unrecognized and therefore unresolved incongruities, or from recall of the comprehension difficulty that was experienced at first exposure.

Mere-exposure effects are the well documented phenomenon wherein positive assessment of stimuli tends to increase with repeated exposures. Positive assessment tends to increase with repeated exposures to stimuli even when people do not actively attend to or actively process the repeated stimuli (Zajonc, 1968; Berlyne, 1970; Jacoby & Kelley, 1987; for a review see Bornstein, 1989; Montoya, Horton, Vevea, Citkowicz &

Lauber, 2017). Mere-exposure effects can potentially account for humour appreciation on repeated exposures due to improvements in processing fluency. Jacoby and Kelley (1987; See also Bornstein & D'Agostino, 1992; 1994) argued that improvements in positive assessment due to mere-exposure effects with repeated exposures can be attributed to improvements in processing fluency for familiar stimuli. Topolinski (2014) analogously argued that fluency of incongruity-resolution should be an important predictor of humour appreciation. Therefore, decreases in the duration of time necessary to understand a pun with repeated exposures should be predictive of humour ratings. Wyer and Collins (1992) countered that mere-exposure effects would be inadequate as an explanation for humour on repetition. Mere-exposure effects should be equally pronounced for all jokes. Therefore, it would not be able to account for why some jokes continue to be funny while others decrease in humour with repetition.

The second possibility proposed by Suls (1972) was that people might be able to work through repeated humorous stimuli in different ways; discovering new incongruities and therefore also being able to reach new humorous resolutions. However, there must be an upper limit to possible interpretations of incongruity. The final explanation for humour on repetition proposed by Suls (1975), from an incongruity-resolution perspective, was that humour from first exposure comes from incongruity-resolution, whereas humour on subsequent exposures could be produced by satisfaction from accurately recognizing and recalling an item. In support of the hypothesis that humour on repetition comes from accurate recall, Suls (1975) conducted an experiment in which undergraduate students were asked to rate and memorize a set of 24 jokes on 3 occasions (participants returned to rate the jokes again after one week and then again after three months). The jokes that were accurately recalled increased in humour, whereas jokes that were only partially recalled were rated as being less humorous. Based on these findings, it could be hypothesized that latent semantic incongruity should be most strongly associated with humour on first exposure, and comprehension difficulty on first exposure should be predictive of humour on all subsequent exposures to a pun.

The comprehension-elaboration model (Wyer & Collins, 1992) also attempted to account for changes in humour appreciation with repeated exposures. In accordance with

incongruity-resolution, humour from a "comprehension" process can be produced from the challenge of explaining humorous incongruities on first exposure to an item. A moderate degree of comprehension difficulty, according to the time and effort necessary to interpret incongruous stimuli, was hypothesized to produce the greatest amount of humour (a hypothesized non-monotonic, inverted-U shaped relation between the duration of time necessary to understand incongruities and humour appreciation). Wyer and Collins predicted that there should also be a second process, beyond interpretation of incongruity, which can also produce humour. Humour appreciation was hypothesized to also come from the linear quantity of post-comprehension time and effort participants are willing to spend elaborating on the humorous aspects of the stimuli. Elaboration was defined as the generation of novel ideas, features, imagery or concepts in association with the humorous aspects of stimuli. Stimuli with a greater potential for the generation novel elaborations should be more likely to have content that is not fully considered on first exposure and thus items with high "elaboration potential" should be more likely to remain humorous over repeated exposures. On the other hand, items with low elaboration potential would produce minimal humour when repeated even once.

3.3 Processing Challenge and Humour

In the current study, comprehension difficulty was further investigated as a predictor of changes in humour appreciation with repeated exposures to written puns. The incongruity-resolution model (Suls, 1972) and comprehension-elaboration model (Wyer & Collins, 1992) both predict that a moderate level of comprehension difficulty should be optimally associated with humour for novel stimuli (neither too difficult nor too hard to understand; that is, a non-linear, inverted-U shaped function). For stimuli such as written puns, which should be relatively easy to understand, greater levels of comprehension difficulty should be positively associated with humour. In contrast, the fluency account of humour appreciation (Topolinski, 2013) predicts that lower comprehension difficulty (faster and easier processing) should always be positively associated with humour. Although it is unclear if comprehension difficulty can continue to play a role in humour appreciation for repeated exposures to a pun, according to the fluency account decreases

in comprehension difficulty with repeated exposures should be predictive of changes in humour ratings.

3.3.1   Comprehension Duration

It is possible that the pun identification task employed in study 1 was too easy or not representative of what would be necessary to fully understand a pun. Participants may have been able to recognize that an item was a pun before adequately understanding the intended humour. In the study reported here, a more direct and representative task that also minimizes distractions from simply enjoying each item was used. The duration of time from the presentation of each pun until they press a button labeled "got it" to indicate that they understand each item (either understood the humour from a pun item or understood that an item was a non-pun control item) was recorded as the "comprehension duration".

3.3.2   Rating Duration

According to the comprehension-elaboration model (Wyer & Collins, 1992), humour from incongruity should also come from the duration of time participants spend elaborating on humorous incongruity after comprehension. Therefore, the duration of time from pressing the "got it" button until participants provided humour ratings and pressed the "next" button to move on to the next pun was also recorded as a variable of interest; referred to as "rating duration". Rating duration, as a measure of the time participants spend considering humour (or re-considering humour ratings for repeated exposures) after comprehension should also be positively associated with humour as a very rough approximation of time spent on post-comprehension "elaboration". The effect of rating duration should be more clearly visible as a significant predictor of humour after the first exposure to a pun.

3.4   Mirth and Cognitive Appraisal

The current study investigates the hypothesis that changes in humour appreciation with repeated exposures may depend on how humour appreciation is measured. Ratings of both "mirth" (the current emotional experience of humour in response to an item) and

"cognitive appraisal" of humour (a more objective overall assessment of the quality of the pun itself) were used to assess humour over repeated exposures. This distinction is based on work by Gavanski (1986), who predicted that repeated exposures to a joke should reduce people's subjective enjoyment without actually changing their evaluation of the objective quality of an item. To test this hypothesis, Gavanski conducted a study in which twenty-five cartoons were each presented to participants up to five times each. Consistent with his hypotheses, participant ratings of mirth (but not cognitive appraisal) significantly decreased with repetition (similar results were obtained by Pistole & Shore, 1979). Gavanski's participants could likely appreciate that the cartoons were objectively still as clever on repetition as they were on first exposure, even if the items did not provoke the same emotional response. Further, Gavanski (1986) argued that without clear instructions to judge humour according to their current subjective emotional experience of humour (mirth), participants tend to respond to humour rating scales according to a more objective "cognitive evaluation" of the quality of a humorous item's overall quality. According to this argument, prior studies that did not make this distinction (as in study 1) may have underestimated decreases in humour with repeated exposures or due to familiarity with a pun from a prior occasion. Recall that in study 1 participants ratings of cleverness, funniness, clarity of understanding, and the extent to which an item is a good example of a pun where all very highly correlated. It could be speculated that these variables each measured a cognitive appraisal of humour, and therefore these dependent variables may have underestimated decreases in humour for familiar stimuli.

3.5 Hypotheses

3.5.1   Humour Ratings and Repetition

The incongruity-resolution model (Suls, 1972) holds that humour comes from the process of comprehending (resolving) incongruities. However, study 1 found that familiarity with a pun from a prior occasion was positively associated with humour ratings. The current study therefore tests the core assumption that humour ratings for puns will on average decrease with repeated exposures.

Gavanski (1986) argued that changes in humour with repetition should depend on how the repeated stimuli is rated. He predicted that the emotional experience of humour "mirth" should decrease with repetition, but the more objective overall assessment of humour, "cognitive appraisal", should not significantly change. That is, repeated stimuli may not present the same emotional experience of humour on repeated exposures, but people should still be able to recognize the overall quality or cleverness of the item.

3.5.2   Latent Semantic Incongruity

If humour is produced by the comprehension of incongruities (as would be predicted by both incongruity resolution and comprehension-elaboration; Suls, 1972; Wyer & Collins, 1992), then the relationship between latent semantic incongruity and participant humour ratings should decrease with repeated exposures (given that the item has already been adequately understood). Latent semantic incongruity should therefore significantly moderate changes in both mirth and cognitive appraisal ratings with repetition. If incongruity from semantic dissimilarity is resolved after first comprehension of a pun, then the association between latent semantic incongruity and humour should decrease with repeated exposures.

3.5.3   Duration Effects

According to predictions from both the incongruity-resolution (Suls, 1972) and comprehension-elaboration models (Wyer & Collins, 1992) a moderate level of comprehension challenge should be associated with humour. There should therefore be a non-linear (inverted-U shaped) relationship between comprehension duration and both mirth and cognitive appraisal ratings. According to the fluency account (Topolinski, 2014), decreases in processing with repetition should work to minimize decreases in humour with repetition. If fluency improves with familiarity, then both comprehension duration and rating duration should decrease with repetition. That is, participants should be able to understand the humour in written puns (comprehension duration) and provide both mirth and cognitive appraisal ratings (rating duration) faster on repeated exposures.

If participants simply recall comprehension difficulty that was experienced on first exposure (as was predicted by Suls, 1975), then comprehension duration on first exposure

should predict humour ratings for all repeated exposures to a pun. According to the comprehension-elaboration model (Wyer & Collins, 1992), the duration of time spent (re)considering and elaborating on the humour in puns should be positively predictive of humour ratings on repeated exposures. As a coarse test of this prediction, rating duration (but not comprehension duration) should therefore be predictive of humour on repeated exposures.

3.6 Method

3.6.1 Participants

In total, 194 participants were recruited at the University of Western Ontario but 13 participants were removed from analyses as outliers (due to low item-total correlation with responses from other participants). Therefore, data from, 181 undergraduate participants (82 females, 99 males; mean participant age = 18.81, $SD = 1.70$) were included in the reported analyses. Participants either spoke English as their first language (171 participants) or they had been speaking English as their primary language for at least 10 years (10 participants; mean years of English experience = 14.00; $SD = 2.78$). Students who participated in study 1 were not eligible to participate in study 2.

3.6.2 Item Types

In total, 230 unique items were selected for the current study: 40 key repetition items, 180 non-repeated filler items, and 10 lure items included to assess participant inattention. Ten lists of items were employed (the number of participants who were assigned to complete each list is provided in Table 4). Each list contained the same set of 180 non-repeated filler items, and the same 10 lure items. The only difference between the 10 lists was that they contained a different set of 4 key repetition items (thus, across the 10 lists, there were 30 repeated pun items and 10 repeated control items in total). Participants were randomly assigned to work through one of 10 item lists.

Table 4: Study 2 participant reliability (according to Cronbach's Alpha) and the number of participants who rated each of the 10 item lists (N).

	N	Mirth α	Cognitive Appraisal α	Comprehension Duration α	Rating Duration α
Non-Repeated Filler Items	181	0.99	0.99	0.96	0.65
Lure Items	181	0.67	0.47	0.71	0.39
Average Reliability for Repetition Items	181	0.86	0.91	0.86	0.61
Repetition Item List 1	12	0.80	0.80	0.89	0.22
Repetition Item List 2	19	0.91	0.93	0.89	0.51
Repetition Item List 3	20	0.82	0.94	0.79	0.54
Repetition Item List 4	19	0.91	0.93	0.88	0.67
Repetition Item List 5	19	0.82	0.90	0.88	0.67
Repetition Item List 6	19	0.76	0.87	0.86	0.63
Repetition Item List 7	19	0.82	0.88	0.91	0.77
Repetition Item List 8	18	0.95	0.96	0.86	0.80
Repetition Item List 9	17	0.87	0.94	0.82	0.50
Repetition Item List 10	19	0.94	0.91	0.82	0.74

Note: Results are reported broken down for each of the four dependent variables and further broken down by list for the repetition items.

3.6.2.1   Repetition Items

Each list contained four "repetition" items that were each repeated five times over the course of the experiment. The 10 lists each contained a different set of 4 repetition items (one homophone, one homograph, one rhyme, one control per list). Across the 10 lists there was a total of 30 repetition pun items (10 homophone, 10 homograph, 10 rhyme) and 10 repeated control items. Repetition items were presented equally spread out over the course of the experiment such that participants would be presented with a repeated item approximately every 40 trials (randomly staggered by plus or minus one to reduce the likelihood that a pattern of repetitions could be noticed). Participant ratings of familiarity from a prior occasion was not used in study 2 to minimize explicit participant attention to the repeated items.

3.6.2.2   Filler Non-Repetition Items

Filler non-repetition items were included to distract participants from the small subset of repetition items. Each list contained the same set of 180 filler non-repetition items that were each presented to participants only once over the course of the experiment (45 homograph based puns, 45 homophone based puns, and 45 rhyme based puns, 45 control items).

3.6.2.3   Lure Items

Ten lure items were included to assess participant inattention. The lure items each described a different disappointing situation that should not be considered humorous (for example, as in the lure item "My parents recently got a divorce. It has been very challenging to deal with").

3.6.2.4   Item Selection

The puns used here were collected from the public submission databases available at punsandjokes.com and punoftheday.com. Puns with racist, sexist, offensive, or excessively aggressive content were not selected to minimize ethical concerns. One hundred and seventy-eight of the 230 items that were selected for study 2 were chosen from those used in study 1. Both repetition items and the non-repetition filler puns were

selected from the items used in study 1 based on funniness ratings (33% from those with low funniness ratings, 34% from those with medium funniness ratings, and 33% from those with high funniness ratings) and with an equal proportion of items based on homophone, homograph, rhyme, and the control items. Control items were constructed from actual puns used in the current study according to the same criterion used in study 1: the semantically ambiguous key word of a pun was replaced with a non-ambiguous keyword that is consistent with the dominant meaning of the sentence. Ten lure items were also included in study 2 to assess participant inattention. The complete list of repetition items, filler items, and lure items used in study 2 is available in Appendix A.

3.6.3   Latent Semantic Incongruity

The same latent semantic incongruity measure that was developed in study 1 was employed again in study 2. Latent semantic analysis provides an estimate of the semantic similarity between two text passages. The current study reverse codes this measure such that results discussed as having "high semantic incongruity" are those with "low semantic similarity". New dictionary definition entries and semantic similarity estimates using LSA were prepared for the 19 new homophone items using the dictionary definition approach described in study 1. Example latent semantic incongruity estimates are available in Appendix B.

3.6.4   Procedure

The study was completed using a custom internet survey platform developed by Dr. Rod Martin. Participants in this study completed the internet-based task on in-lab computers at the University of Western Ontario. Prior to starting the experiment, participants were asked to provide general demographics: age, gender, whether English was their first language, and if English was their second language, the number of years speaking English as their primary language. Participants were first presented with each item in an otherwise empty screen and were instructed to click a "got it" button once they have read and understood the item. In the case of control items, they were instructed that they should push the "got it" button once they sufficiently understood that the given item was not actually a pun. The duration of time between the initial presentation of each item and

clicking the "got it" button was recorded as "comprehension duration". Next, participants were asked to rate the extent to which each item causes them to experience the emotional state of mirth (defined as the *emotional* experience of amusement, fun, up to hilarity; typically accompanied by the urge to grin, smile or laugh) on a 7-point Likert-type scale (ranging from 0 not at all, to 6, strong experience of mirth). Participants were also asked to rate their "cognitive appraisal" of each item according to the extent to which they *think* that each item is an effective example of humour (defined as being clear, clever, and an overall good example of humour; something that you think other people would enjoy) on a 7-point likert type scale ranging from 0 not at all, to 6 very effective. These definitions of mirth and cognitive appraisal of humour were provided on all screens of the experiment. To address ethical concerns, participants were also allowed to select "no answer" for mirth or cognitive appraisal ratings if they did not feel comfortable assessing an item. The "no answer" option was selected only 202 of 38010 cases (0.5% of the time) for ratings of mirth and 205 of 38010 cases (0.5% of time for their cognitive assessment). The duration of time between pressing the "got it" button for each item, providing mirth and cognitive appraisal ratings, and pressing the "next" button was recorded as "rating duration".

3.6.5 Data Preparation

3.6.5.1 Missing Data

Duration values with a greater duration than 25000 ms (extreme values selected approximately based on three standard deviations from the mean duration scores in the raw data) were discarded and replaced with the mean over participants for that item for both comprehension duration (138 replaced; 0.4% of 38010 cases) and rating duration (160 replaced; 0.4% of 38010 cases). Missing data due to participant selection of the "no answer" option was also replaced with the mean over participants for mirth (167 replaced; 0.4% of 38010 cases) and cognitive appraisal ratings (172 replaced; 0.5% of 38010 cases).

Inferential statistics were conducted on data that was averaged across items. Given that participants were assigned to rate one of ten item lists, it was not possible to calculate

averages over participants. Averages over items were prepared for comprehension duration and rating duration (both reported in milliseconds), and for the Likert-type ratings for mirth and for cognitive assessment.

3.6.5.2   Internal Reliability

Participant item-total correlations were examined to assess the extent to which participant responses are consistent with average participant responses. Data from participants with low item-total correlations according to any one scale (lower than .2) were removed from analysis as atypical outliers. Using this criterion, data from 13 participants were removed from final analyses, leaving 181 participants included in all subsequent reported analyses.

The internal consistency of participants in study 2 was assessed using Cronbach's alpha reliability analyses (with participant responses treated as items on a scale). Participant reliability was assessed for ratings of mirth, cognitive appraisal, comprehension duration, and rating duration. Participant reliability estimates were assessed separately for lure items, non-repetition items, and repetition items. Reliability for repetition items was assessed in separate analyses for each of the 10 lists; average reliability for repetition items over the 10 lists is reported. Refer back to Table 4 for the complete list of reliability scores across the 10 item lists for each dependent variable.

Across scales, participant reliability for lure items was low (ranging from .39 to .71), indicating that participants less clearly agreed on their assessments of these items. Otherwise, with 181 participants included, participant reliability as assessed using Cronbach's alpha was high for ratings of both mirth (average repetition item reliability α = .86, non-repeated items reliability α = .99) and cognitive appraisal of humour (average repetition item reliability α = .91, non-repeated items reliability α = .99). Participant reliability for comprehension duration variables was also high (average repetition item reliability α = .86, non-repeated items reliability α = .96). Participants were less clearly consistent (although reliability was still satisfactory), for rating duration (average repetition item reliability α = .61, non-repeated items reliability α = .65). The relatively lower reliability for rating duration indicates that participants were more variable in the

quantity of time they took in considering and selecting mirth and cognitive appraisal ratings.

For lure items, Cronbach's alpha estimates were comparatively low: Cronbach's alpha was .67 for ratings of mirth, .47 for cognitive assessment, and .71 for response duration to "get" the item. The ten lure items were originally included in study 2 with the intention of using them as additional cues to participant inattentiveness. Recall that the lure items were non-puns that describe a disappointing situation (for example, "In the last round of downsizing most of my co-workers were fired. I will probably be next"). However, it would appear that the lure items were ineffective as a tool to detect participant inattention given that a subset of participants found these items humorous (27 of 181 participants had cognitive appraisal humour ratings for these items > 1). Upon further examination, participants who found the lure items humorous responded in a pattern that otherwise appeared to be reliable and consistent with other participants, with strong item-total correlations and normal response variability, so their responses were not removed from further analyses.

3.7 Results

3.7.1   Mirth and Cognitive Appraisal

Mean and standard deviations for the complete set of dependent variables (mirth, cognitive appraisal ratings of humour, comprehension duration, and rating duration) broken down by item type and further separated by each of 5 exposures is presented in Table 5. Overall, the number of exposures to the puns was negatively correlated with ratings of both mirth $r(148) = -.40$, $p < .001$, and cognitive appraisal $r(148) = -.29$, $p < .001$. This finding confirms that, overall, both humour ratings significantly decreased with repeated exposures. The number of exposures to a pun was also associated with significantly faster responses according to both comprehension duration $r(148) = -.71$, $p < .001$ and rating duration $r(148) = -.55$, $p < .001$. That is, with repeated exposures both comprehension and rating durations became shorter. The overall correlation between dependent variables for the repetition pun items is presented in Table 6. Mirth ratings were significantly lower than cognitive appraisal ratings across all five exposures to

Table 5: Mean and standard deviation over items for pun and control items for each of the four study 2 dependent variables.

Dependent Variable	Item Type	N	Pun items Mean (*SD*)	N	Control Items Mean (*SD*)
Latent Semantic Incongruity	Non-Repeated Filler	135	.49 (.13)		
	Repetition Items	150	.48 (.12)		
Mirth	Non-repeated Filler	135	2.63 (0.57)	45	0.82 (0.4)
	Repetition items	150	2.26 (0.76)	50	0.5 (0.24)
	1st exposure	30	2.52 (1.1)	10	0.69 (0.23)
	2nd exposure	30	2.25 (0.97)	10	0.48 (0.25)
	3rd exposure	30	1.99 (0.95)	10	0.43 (0.22)
	4th Exposure	30	1.87 (0.83)	10	0.45 (0.22)
	5th Exposure	30	1.73 (0.71)	10	0.45 (0.23)
Cognitive Appraisal	Non-repeated Filler	135	2.89 (0.56)	45	0.88 (0.44)
	Repetition items	150	2.71 (0.8)	50	0.66 (0.42)
	1st exposure	30	2.83 (1.1)	10	0.87 (0.4)
	2nd exposure	30	2.58 (1.08)	10	0.62 (0.45)
	3rd exposure	30	2.39 (1.08)	10	0.57 (0.37)
	4th Exposure	30	2.34 (1.04)	10	0.62 (0.44)
	5th Exposure	30	2.28 (1.02)	10	0.58 (0.44)
Comprehension Duration	Non-repeated Filler	135	4182.2 (1055.98)	45	4772.74 (1547.52)
	Repetition items	150	2976.38 (1280.21)	50	3742.34 (2918.08)
	1st exposure	30	4932.23 (1213.93)	10	5966.49 (1707.03)
	2nd exposure	30	3140.7 (802.4)	10	5252.24 (5382.18)
	3rd exposure	30	2499.58 (520.09)	10	2721.75 (677.55)
	4th Exposure	30	2341.13 (732.52)	10	2380.7 (473.54)
	5th Exposure	30	2219.56 (622.05)	10	2381.72 (699.03)
Rating Duration	Repetition items	150	4633.53 (943.01)	50	4241.99 (1039.05)
	1st exposure	30	5526.52 (787.14)	10	5600.49 (1198.49)
	2nd exposure	30	4793.28 (531.71)	10	4630.72 (675.73)
	3rd exposure	30	4193.96 (437.2)	10	3781.83 (333.7)
	4th Exposure	30	4388.47 (1186.57)	10	3660.38 (450.87)
	5th Exposure	30	3955.07 (760.67)	10	3504.18 (368.2)

Note: Averages are reported broken down by item type (non-repeated filler item, lure items, and repetition items). Averages for repetition items are further broken down by number of exposures: from the first exposure (1) to the fifth exposure (5) to a pun. Mirth and cognitive appraisal were on scales from 0 (not at all) to 6 (extremely). Comprehension Duration and rating duration were recorded in milliseconds.

Table 6: Overall correlation between Study 2 dependent variables for the repeated items.

	1	2	3	4	5	6
1. Latent Semantic Incongruity	-	-.02	0.32**	0.42**	-.17*	-.02
2. #Of Exposures	-	-	-0.40*	-0.29*	-0.71**	-0.55**
3. Mirth	-	-0.30*	-	0.89**	0.25**	0.18*
4. Cognitive Appraisal	-	-0.20	0.88**	-	0.16*	0.10
5. Comprehension Duration	-	-0.49**	0.32*	0.26	-	0.57**
6. Rating Duration	-	-0.72**	0.54**	0.50**	0.39*	-

Note: Correlations amongst repeated pun items (30 items each presented five times; N = 150) are presented above the diagonal. Intercorrelations for repeated control items (N=10 items each presented 5 times; N = 50) are presented below the diagonal. Correlations marked by (*) were statistically significant at $p < .05$, Correlations marked by (**) were statistically significant at $p < .01$.

written puns (according to paired sample t-tests; all comparisons significant at $p < .001$; see Table 7).

3.7.2 Latent Semantic Incongruity

3.7.2.1 Cognitive Appraisal

Stepwise multiple regression analyses were conducted to investigate latent semantic incongruity as a moderator of changes in cognitive appraisal ratings of humour with repeated exposures. At the first step, there was a significant effect of the number of exposures to a pun on participant cognitive appraisal ratings $\beta = -.17$, $p < .001$; $R^2 = .09$, $F(1, 148) = 14.69$, $p < .001$. At the second step, there was a significant effect of the number of both exposures to a pun $\beta = -.17$, $p < .001$ and latent semantic incongruity $\beta = 2.68$, $p < .001$; $R^2 = .25$, F change $(1, 147) = 30.62$, $p < .001$. At the final step, there was a significant main effect of latent semantic incongruity $\beta = 2.80$, $p < .05$, but no significant main effect of the number of exposures $\beta = -0.15$, ns, and there was no statistically significant moderating interaction term $\beta = -0.04$, ns; $R^2 = .25$, F change $(1, 146) = 0.01$, ns. For ratings of cognitive appraisal, latent semantic incongruity was significantly associated with cognitive appraisal ratings of humour at all exposures to a pun (see Table 8).

3.7.2.2 Mirth

Stepwise multiple regression analyses were conducted to investigate latent semantic incongruity as a moderator of changes in mirth with repeated exposures. At the first step, there was a significant effect of the number of exposures to a pun on participant mirth ratings $\beta = -.22$, $p < .001$; $R^2 = .17$, $F(1, 148) = 29.78$, $p < .001$. At the second step, there was a significant effect of the number of exposures to a pun $\beta = -.22$, $p < .001$, and latent semantic incongruity $\beta = 1.90$, $p < .001$; $R^2 = .26$, F change $(1, 147) = 17.48$, $p < .001$. At the final step, there was no main effect of the number of exposures to a pun $\beta = .10$, ns, but there was a significant main effect of latent semantic incongruity $\beta = 3.95$, $p < .001$ and the moderating interaction term was statistically significant $\beta = -.67$, $p < .05$; $R^2 = .28$, F change $(1, 146) = 4.62$, $p < .05$. The changes in mirth with repeated exposures to the puns was significantly influenced by semantic incongruity.

Table 7: Comparison between average mirth and cognitive appraisal at each of five exposures to repeated pun items.

Exposure	N	Mirth Mean (*SD*)	Cognitive Appraisal Mean (*SD*)	Mirth Vs. Cognitive Appraisal
Non repeated filler	135	2.63 (0.57)	2.89 (0.56)	$t(134) = 23.29 \ p < .001$
1	30	2.52 (1.10)	2.83 (1.10)	$t(29) = 6.61 \ p < .001$
2	30	2.25 (0.97)	2.58 (1.08)	$t(29) = 6.10 \ p < .001$
3	30	1.99 (0.95)	2.39 (1.08)	$t(29) = 7.02 \ p < .001$
4	30	1.87 (0.83)	2.34 (1.04)	$t(29) = 6.25 \ p < .001$
5	30	1.73 (0.71)	2.28 (1.02)	$t(29) = 5.65 \ p < .001$

Note: Mean and standard deviation (SD) for each study 2 dependent variable over five exposures to a written pun and paired sample *t*-tests (matched by item; n=30) of the difference between mirth and cognitive appraisal . All five comparisons were significant at $p < .001$.

Table 8: Correlation between latent semantic incongruity and study 2 dependent variables over five exposures for repetition pun items.

						Latent Semantic Incongruity					
Mirth 1	0.47**	Cognitive Appraisal 1	0.45*	Comprehension Duration 1	-0.39*	Rating Duration 1	0.19
Mirth 2	0.44*	Cognitive Appraisal 2	0.45*	Comprehension Duration 2	-0.2	Rating Duration 2	0.14
Mirth 3	0.42*	Cognitive Appraisal 3	0.48**	Comprehension Duration 3	-0.27	Rating Duration 3	0.21
Mirth 4	0.36	Cognitive Appraisal 4	0.47**	Comprehension Duration 4	-0.29	Rating Duration 4	-0.11
Mirth 5	0.24	Cognitive Appraisal 5	0.48**	Comprehension Duration 5	-0.24	Rating Duration 5	0.2

Note: Correlations marked by (*) were statistically significant at $p < .05$, Correlations marked by (**) were statistically significant at $p < .01$.

Post hoc analysis of slopes indicated that puns with high latent semantic incongruity (set as .8; approximately two standard deviations above the mean) significantly decreased in humour ratings between the first and fifth exposure $t(149) = -3.94$, $p < .001$. In contrast, for puns with low latent semantic incongruity (set as .2; approximately 2 SD below the mean) mirth ratings did not significantly change with repetition. The significant moderating interaction of latent semantic incongruity on changes in mirth with repetition is illustrated in Figure 5.

Correlation analyses were also conducted to further explore the moderating function of latent semantic incongruity for mirth ratings over repeated exposures to puns. For the 30 repetition pun items, latent semantic incongruity was significantly correlated with mirth at the first exposure $r(28) = .47$, $p < .01$, second exposure $r(28) = .44$, $p < .05$, and third exposure to a pun $r(28) = .42$, $p < .05$. Latent semantic incongruity was not significantly associated with mirth at the fourth exposure $r(28) = .36$, $p < .10$ (although trending), or at the fifth exposure to a pun $r(28) = .24$, ns. For the complete set of correlations between latent semantic incongruity and mirth broken down by number of exposures, refer to Table 8.

3.7.3   Duration Variables

Comprehension duration and rating duration did not significantly moderate changes in mirth or cognitive appraisal with repetition. There was also no statistically significant non-linear effect of comprehension duration on participant humour ratings (for mirth or cognitive appraisal) for the repeated pun items or for the non-repeated filler items. The association between comprehension duration, rating duration and humour was therefore described in greater detail according to bivariate correlations between comprehension duration and humour at each exposure to the repeated written puns.

3.7.3.1   Comprehension Duration

Latent semantic incongruity was significantly correlated with shorter comprehension durations, but only at first exposure $r(28) = -.39$, $p < .05$ (see Table 8). Rather than posing a comprehension challenge, puns with greater semantic incongruity were easier to understand. Comprehension duration at first exposure to a pun was
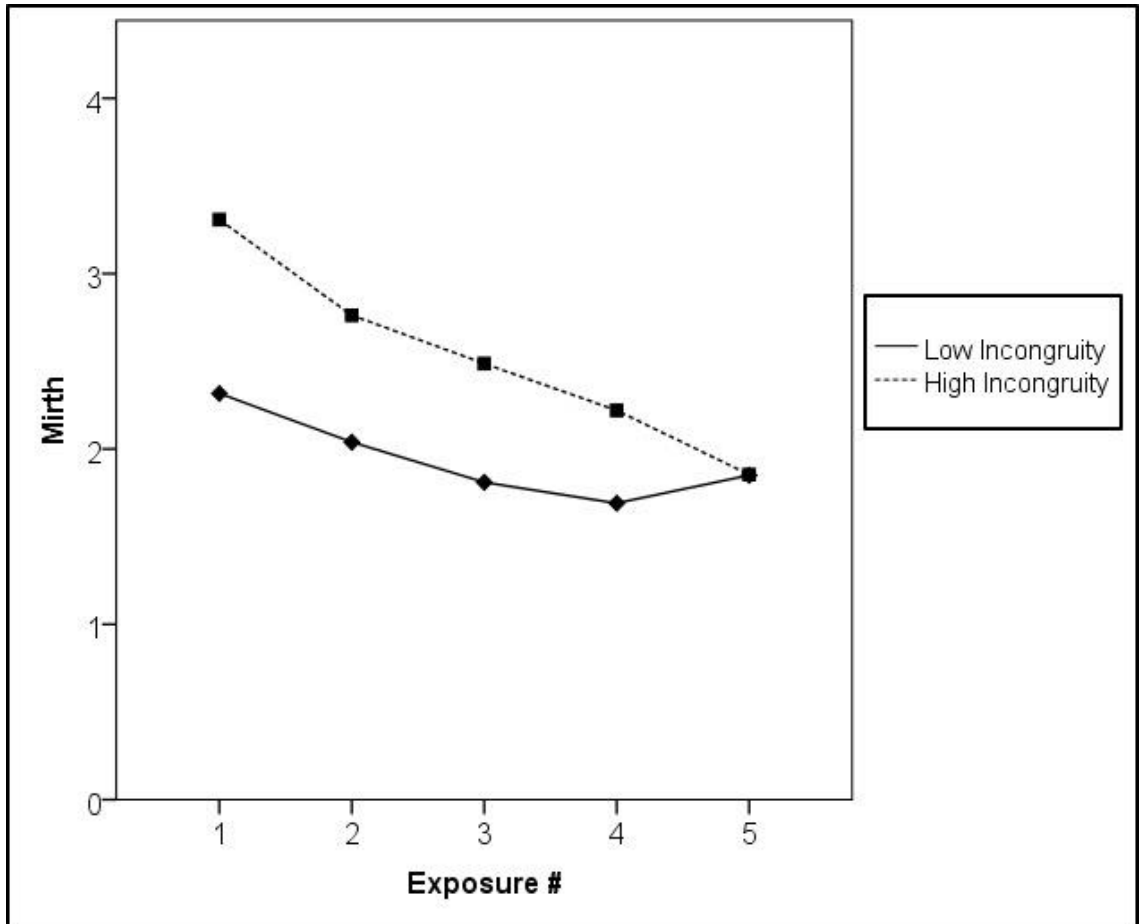
Figure 5: Moderating effect of latent semantic incongruity on changes in mirth with repetition in study 2.

Puns with high latent semantic incongruity significantly decreased in mirth with repetition. There were no significant changes in mirth for puns with low latent semantic incongruity.

significantly negatively associated with both mirth ratings and cognitive appraisal ratings at all exposures. That is, shorter durations of time necessary to "get" the pun at first exposure was not only associated with humour at first exposure, but also with humour at each of the four subsequent repetitions. Comprehension duration at second, third or fourth exposure was not significantly associated with participant mirth or cognitive appraisal humour ratings at the respective exposures. The complete set of bivariate correlations between comprehension duration and the dependent variables over 5 exposures to pun items is presented in Table 9. Note the exceptions to the aforementioned pattern: at exposure 1 the correlation between cognitive appraisal and comprehension duration was only trending towards statistical significance $r(28) = -.32, p < .10$). Comprehension duration at fifth exposure was again significantly associated with fifth exposure mirth $r(28) = -.45, p < .05$ and fifth exposure cognitive appraisal ratings $r(28) = -.48, p < .01$.

3.7.3.2   Rating Duration

Rating duration was trending towards a significant positive association with humour at the second exposure [albeit only trending to statistical significance; mirth $r(28) = .33, p < .10$ and cognitive appraisal humour ratings $r(28) = .34, p < .10$] and at the third exposure [for both mirth $r(28) = .51, p < .01$ and cognitive appraisal ratings $r(28) = .48, p < .01$]. That is, longer durations of time considering whether a pun item was humorous on second and third exposures to an item were positively associated with humour ratings. There were no further statistically significant bivariate associations over the number of exposures between rating duration and participant humour ratings. Bivariate correlations between ratings duration and humour appreciation variables (mirth, and cognitive appraisal) are presented broken down by number of exposures in Table 10.

3.8   Discussion

The primary goal of the current study was to provide an empirical test of how humour is processed and appreciated over repeated exposures to written puns. The incongruity-resolution model (Suls, 1972) emphasized that incongruity must be adequately explained

Table 9: Correlation between comprehension duration and study 2 dependent variables over five exposures to repetition pun items.

Exposures	Comprehension Duration				
	1	2	3	4	5
Mirth 1	-0.36*	-0.21	-0.12	-0.15	-0.46*
Mirth 2	-0.40*	-0.14	-0.15	-0.16	-0.48*
Mirth 3	-0.41*	-0.12	-0.19	-0.17	-0.47*
Mirth 4	-0.36*	-0.14	-0.20	-0.14	-0.44*
Mirth 5	-0.40*	-0.20	-0.19	-0.18	-0.45*
Cognitive Appraisal 1	-0.32	-0.22	-0.13	-0.15	-0.52**
Cognitive Appraisal 2	-0.39*	-0.17	-0.10	-0.10	-0.48**
Cognitive Appraisal 3	-0.41*	-0.14	-0.20	-0.14	-0.50**
Cognitive Appraisal 4	-0.42*	-0.18	-0.19	-0.19	-0.50**
Cognitive Appraisal 5	-0.51**	-0.25	-0.16	-0.23	-0.48**
Rating Duration 1	0.32	-0.04	-0.03	0.00	-0.11
Rating Duration 2	0.18	-0.03	0.31	0.25	-0.10
Rating Duration 3	0.10	0.06	-0.11	-0.33	-0.50**
Rating Duration 4	0.14	-0.11	0.08	0.40*	0.05
Rating Duration 5	0.05	-0.06	0.04	0.18	-0.14

Note: Correlations marked by (*) were statistically significant at $p < .05$, correlations marked by (**) were statistically significant at $p < .01$.

Table 10: Correlation between rating duration and humour appreciation variables (mirth and cognitive appraisal) over five exposures to repetition pun items.

	Rating Duration				
	1	2	3	4	5
Mirth 1	0.11	0.38*	0.53*	-0.03	0.06
Mirth 2	0.07	0.33	0.51**	-0.02	0.07
Mirth 3	0.11	0.31	0.51**	-0.07	0.03
Mirth 4	0.15	0.31	0.50**	-0.08	0.05
Mirth 5	0.09	0.26	0.52**	-0.09	-0.01
Cognitive Appraisal 1	0.13	0.41*	0.53**	0.00	0.08
Cognitive Appraisal 2	0.06	0.34	0.43*	0.08	0.13
Cognitive Appraisal 3	0.11	0.35	0.48**	0.00	0.06
Cognitive Appraisal 4	0.10	0.33	0.47**	-0.05	0.04
Cognitive Appraisal 5	0.03	0.29	0.45*	-0.06	0.05

Note: Correlations marked by (*) were statistically significant at $p < .05$, correlations marked by (**) were statistically significant at $p < .01$.

(resolved) for it to be experienced as humorous. Resolution of incongruity was operationalized and tested as the measurable changes in humour from semantic incongruity over repeated exposures to a pun. In support of this operationalization and consistent with incongruity-resolution based predictions, the association between latent semantic incongruity and participant humour ratings significantly decreased over repeated exposures to a pun according ratings of mirth (but not for cognitive appraisal ratings). The fluency of participant comprehension of the pun items, rather than greater comprehension difficulty or an inverted-u shaped function, was found to be important to humour. Further, comprehension duration on first exposure was predictive of humour at subsequent exposures (consistent with the hypothesis that humour on repetition may come from successful recall of comprehension at a prior exposure; Suls, 1975). Finally, consistent with the comprehension-elaboration model (Wyer & Collins, 1992), longer durations of post-comprehension time spent considering humour ratings (according to rating duration; as a very loose proxy for elaboration) at second and third exposure were positively associated with humour appreciation according to ratings of both mirth and cognitive appraisal. The implications of these findings are discussed below.

3.8.1   Mirth and Cognitive Appraisal

Decreases in humour appreciation for puns with high semantic incongruity proved to be dependent on how humour was measured. Consistent with Gavanski's (1986) findings, latent semantic incongruity moderated decreases in mirth, but not cognitive appraisal ratings of humour over five exposures to puns. Mirth from puns with high semantic incongruity, but not for puns with low semantic incongruity, significantly decreased with repetition to the point that there was no longer a significant effect of semantic incongruity on mirth at the fourth or fifth exposure. That is, the most semantically incongruous puns produced less mirth with repetition whereas participant assessment of less semantically incongruous puns did not change. This finding indicates that the experience of mirth may actually only decrease with repetition for the most incongruous (or potentially just the most humorous) stimuli. The cognitive appraisal ratings of humour for puns with high semantic incongruity was always significantly more humorous than puns with low latent semantic incongruity. Thus, as the stimuli became increasingly familiar, semantically

incongruous puns were less likely to evoke the emotional experience of mirth, but they continued to be seen as good examples of humour according to ratings of cognitive appraisal. Across all five exposures to the repeated puns, the more objective cognitive appraisal ratings of the overall quality of the humour in the puns were always significantly higher than ratings of mirth (a more subjective and immediate emotional experience of humour in response to the puns). Participants appear to have appreciated that their current experience of humour is distinct (and lower) than the more objective quality of an item.

Prior studies which evaluated the effectiveness of stimuli, rather than participant current mirth reactions may have underestimated changes in humour ratings with repeated exposures. For example, Zhang and Zinkhan (1991) did not distinguish between mirth and cognitive appraisal and found that "perceived humour" from advertisements did not significantly change within three repetitions (analogous to Gavanski's results from ratings of cognitive appraisal). Suls (1975) asked participants to assess how funny his joke items were over three repeated exposures. He found that funniness of the items did not significantly change with repetition of a week and then again at 3 months later (except for only partially remembered jokes at second exposure). Study 1 participants were asked to rate the quality of the pun, rather than their emotional experience, and therefore may have overestimated the positive relationship between humour and familiarity ratings.

3.8.2   Comprehension Fluency and Humour

There were three important findings regarding comprehension duration and humour: comprehension duration did not moderate changes in humour ratings with repetition, latent semantic incongruity was associated with shorter comprehension durations, and shorter comprehension durations at first exposure were positively associated with humour ratings across all five exposures to a pun. These findings support several predictions from incongruity-resolution, but only if accommodations are made to the extant theory taking into account the role of fluency in humour appreciation.

Comprehension duration (the time necessary to "get" a pun) significantly decreased with repeated exposures. However, improvements in fluency did not significantly moderate changes in perceived humour with repeated exposures. That is, the puns did not become more (or less) humorous with repetition because they became easier to process. It was hypothesized that improvements in processing fluency could account for preserved humour on repeated exposures. However, the observed decreases in comprehension and rating durations across repeated exposures to puns could not account for persistent humour appreciation for familiar stimuli.

In the current study, latent semantic incongruity was associated with shorter comprehension durations. Prior models assume that incongruity is humorous because it presents a challenge for comprehension (Suls, 1972; Wyer & Collins, 1992; Ritchie, 1999, 2004; Forabosco, 1992). Stimuli with relatively higher levels of incongruity were assumed to be more difficult to accurately comprehend (requiring a greater level of complexity in problem solving; Suls, 1972). Instead what was found here was that semantic incongruity was correlated with shorter comprehension durations. This finding is supportive of the fluency account. The association between shorter comprehension durations and humour ratings is correlational and therefore directionality of this effect is unclear: humorous semantic incongruity may facilitate information processing speed or fluency may produce humour from semantic incongruity.

Shorter comprehension duration at first exposure was also predictive of humour ratings across all exposures (both at first exposure and over all subsequent repeated exposures). With accommodations for the importance of fluency, this finding is supportive of Suls' (1975) account of humour on repetition. Humour on the first exposure was hypothesized to come from comprehension challenge, whereas humour experienced on repetition should come from the satisfaction associated with successfully recognizing or recalling an item. Participants were likely able to either recall the ease at which the puns were understood, or upon recognition deduce how difficult it must have been to understand the item. These results are also mostly inconsistent with Suls' (1972) hypothesis that puns can be humorous on repetition due to re-interpretation. There was no evidence that re-interpretation could account for humour appreciation on the second to fourth exposures to

the pun items which limits the usefulness of this explanation. The notable exception was for comprehension duration on the fifth and final exposure to the repeated puns.

3.8.3   Rating Duration and Comprehension-Elaboration

In contrast to the importance of shorter comprehension durations for humour appreciation, *longer* durations of time spent providing humour ratings (rating durations) at second and third exposure to a pun were positively associated with humour.[6] Although this finding is based on a very coarse operationalization of response durations, it is inconsistent with a strong interpretation of Topolinski's fluency account which would predict that shorter processing durations always should produce more humour.

Shorter comprehension durations at first exposures and longer rating durations at second and third exposures were positively associated with humour. According to incongruity-resolution theory (Suls, 1972), participants may have been using this time while providing humour ratings to re-interpret the incongruities of the pun. Latent semantic incongruity may have been correlated with mirth until all possible interpretations of incongruities have been considered. However, if this were the case, it should have been expected that re-interpretation would take place again during the "comprehension duration" period.

Wyer and Collins (1992) predicted that longer durations of time spent on post-comprehension "elaboration" should be positively associated with humour. Elaboration was defined as the conscious generation of novel thoughts and features beyond what is necessary for comprehension. According to this argument, repeated stimuli can remain humorous so long as people are able to generate novel humorous elaborations. This explanation is consistent with current findings if one were to argue that participants were able to generate novel elaborations on second and third exposures to the repeated puns until the potential for novel elaborations was exhausted at the fourth and fifth exposures.

---

[6] Second exposure longer rating durations was only trending towards a significant correlation with second exposure mirth [$r(30) = .33; p < .08$] and cognitive appraisal of humour [$r(30) = .34. p < .06$].

Latent semantic incongruity was significantly correlated with mirth ratings only for the same first three exposures to a written pun. I propose that these results suggest that latent semantic incongruity may be a meaningful indicator of the potential of a pun for the generation of novel elaboration. High semantic incongruity items should have had a greater quantity of novel related concepts that could come in to play. Consistent with this hypothesis, Utsumi (2005), found that participants were able to generate a greater quantity of emergent features for metaphors with lower levels of semantic similarity (and therefore greater semantic incongruity) between referenced concepts. In study 3, this elaboration potential hypothesis will be evaluated.

3.8.4   Evaluation of Mere-Exposure Effects in Repeated Puns

Mere-exposure effects are the well documented phenomenon wherein positive assessment of stimuli tends to increase with repeated exposures. Suls (1972) hypothesized that mere-exposure effects could account for persistent humour appreciation over repetition. To evaluate this argument, current findings will be evaluated according to Jacoby et al.'s explanation of mere-exposure as fluency effects (Jacoby & Kelley, 1987; Jacoby, Toth, Lindsay & Debner, 1992), and according to Berlyne's (1970; see also Stang, 1973) two-factor model of mere-exposure (for a comprehensive discussion of mere-exposure accounts see Bornstein, 1989; Montoya, Horton, Vevea, Citkowicz & Lauber, 2017).

Topolinski's (2014) fluency account of humour was developed out of Jacoby et al.'s fluency account of mere-exposure effects (see Jacoby & Kelley, 1987; Jacoby, Toth, Lindsay & Debner, 1992). Jacoby et al. hypothesized that increased positive assessment with repetition was due to positive feelings associated with fluency caused by a reduction of uncertainty or from feelings of accomplishment from recognition of prior successful comprehension. Topolinski (2014) demonstrated that priming material from the punchline of a joke (1 minute or 15 minutes prior to hearing a joke; but not immediately prior to reading a joke) improved funniness ratings. Topolinski also demonstrated that jokes were rated as being significantly less humorous when presented in a difficult to read font (as opposed to a standard font). However, Topolinski did not investigate whether improvements in processing speed were predictive of changes in humour. The

current finding that decreases in comprehension duration with repeated exposures to puns was <u>not</u> significantly associated with changes in humour ratings is not supportive of a strong interpretation of the fluency account of humour (that rapid processing is always beneficial to humour). Further, longer durations of time spent providing humour ratings at second and third exposure were actually positively associated with humour.

It is proposed in Berlyne's (1970; see also Stang, 1973) two-factor model of mere-exposure effects that improvements in fluency and familiarity with repetition should increase positive assessment of stimuli, but a second process of satiation and boredom should also be at play (which would limit mere-exposure effects). Within a single context (as was the case here) it should then follow that repetition should decrease humour due to boredom and increase a cognitive satiation with the ideas at play. With sufficient time between repetitions, satiation and boredom with an item will decrease and positive assessment will increase due improvements in fluency (for supportive evidence of the two-factor model from a mere-exposure perspective see also Bornstein, Kale & Cornell, 1990). Consistent with predictions from Berlyne's two-factor mere-exposure model, Forabosco (1994) obtained evidence of semantic satiation without direct repetition of content. When Forabosco's participants rated a series of distinct jokes that shared similar semantic content one after another (with a similar topic or themes), the jokes became less humorous. There were larger decreases in humour ratings for jokes in a sequence that were more strongly semantically similar to each other. That is, even though they were new jokes, participants got bored of the similar semantic content. Although Topolinski (2014) demonstrated that priming material from the punchline of a joke increased humour ratings, there was no effect of priming when it took place immediately prior to reading the joke, which suggests that there may have been a consistent satiation or boredom effect. Also consistent with this hypothesis is Goldstein's (1970b; as cited in Suls, 1975) finding that decreases in humour ratings with repeated exposures to humorous cartoons were significantly lower for participants in an aroused state (and who were therefore less likely to be bored). Berlyne's (1970) two-factor model of mere-exposure effects appears to be able to effectively account for current findings: it can potentially reconcile the study 1 finding that familiarity from an unspecified prior occasion was a positive associate of humour (because levels of boredom with the content would have decreased from

exposure in the prior context) with the study 2 finding that humour ratings decrease with repeated exposures to a pun within a single context.

3.8.5   Summary

Consistent with incongruity-resolution based predictions (Suls, 1972), novelty appears to be important to humour. Ratings of mirth (but not cognitive appraisal) significantly decreased for puns with high semantic incongruity (but not for puns with low latent semantic incongruity) across repeated exposures to puns. There was no evidence for the hypothesized non-linear (inverted-u) shaped relationship between comprehension duration and humour. Shorter comprehension duration at first exposure to a pun was associated with humour at both first exposure and for all repetitions. This finding is consistent with both Topolinski's (2014) fluency account, and Suls' (1975) argument that humour on repetition may come from recall of the degree of comprehension challenge that was initially posed by familiar stimuli at first exposure. In contrast to predictions from incongruity-resolution, latent semantic incongruity was associated with shorter comprehension durations. That is, rather than presenting a challenge for comprehension, semantically incongruous puns appear to be easier to understand. It was problematic for the fluency account that reductions in comprehension duration with repetition were not predictive of changes in humour appreciation and that longer durations of time providing humour ratings at second and third exposure were positively associated with humour. The positive association between longer rating durations (at second and third exposure) and humour was interpreted as being consistent with either Suls' (1972) hypothesis that the incongruities were either re-interpreted on second and third exposure, or with Wyer and Collins (1992) hypothesis that post-comprehension elaboration should also be predictive of humour. Participants may have been able to discover novel elaborations during the post-comprehension rating duration. It was further hypothesized that latent semantic incongruity may be a predictor of the potential of a written pun for the generation of novel elaboration. This hypothesis is further investigated in study 3.

Chapter 4

4 Study 3: Comprehension-Elaboration and Humour From Familiar Puns

The primary purpose of this study was to provide an empirical test of the hypothesized function of "elaboration" in humour appreciation. Studies 1 and 2 evaluated empirical predictions regarding the function of incongruity and resolution in humour appreciation. The lingering problem is that incongruity-resolution cannot clearly account for humour from familiar (study 1) or repeated (study 2) stimuli. Wyer and Collins (1992) argued that, after adequate comprehension of incongruities, humour can also be produced by the quantity of cognitive elaboration on humorous stimuli and its implications. Elaboration was defined as the conscious generation of novel thoughts and features in association with humorous stimuli beyond what is necessary for comprehension. They hypothesized that repeated humorous stimuli can remain humorous until people have considered all easily accessible elaborations that come to mind in association with the stimuli.

In the current study, elaboration was operationalized according to participant performance on a free association concept cueing task. Participants were asked to list as many words as possible that come to mind when thinking of each of the two possible implied concepts in written puns. Elaboration quantity, the number of words participants provided in association with the implied concepts from each pun and the duration of time participants were willing to spend on the task for each pun (elaboration duration) were recorded as variables of interest. To test the hypothesis that elaboration after viewing and comprehending an item is more productive for humour appreciation than elaboration prior to viewing the referent pun, participants were assigned to complete the elaboration task with either the pun present for the entire task, or to an experimental condition in which they were asked to elaborate on associated concepts prior to having viewed (and comprehended) the referent pun.

4.1 Comprehension-Elaboration

Recall that Wyer and Collins (1992) predicted that potentially humorous stimuli, such as written puns, are initially interpreted according to the most salient available set of

concepts and schemata. When incongruities occur that cannot be explained according to the initial set of concepts, the stimuli must be re-interpreted according to an application of an additional set of concepts. The quantity of humour produced from this re-interpretation process should depend on the time and effort necessary to correctly re-interpret incongruous stimuli according to a non-linear inverted- U shaped function (for consistent evidence, see Zigler, Levine & Gould, 1966; 1967; McGhee, 1976). The hypothesized "comprehension" process is functionally analogous to what would be predicted from the incongruity-resolution model (as described by Suls, 1972).

Wyer and Collins (1992) also predict that a second "elaboration" process can produce humour. Cognitive elaboration was hypothesized to take place only after the comprehension process is complete. That is, adequate comprehension should be important to humour produced from elaboration. They predicted that there should be a linear positive association between the quantity of cognitive elaboration and humour appreciation. It was hypothesized that humour can be produced from both the challenge of incongruity comprehension and from elaboration. It was further argued that stimuli should remain humorous on repeated exposures to stimuli so long as novel elaborations can come to mind (at least until all easily accessible novel elaborations have been considered). Items with greater potential for the generation of novel elaborations should remain humorous across a greater number of repeated exposures.

In study 2, for second and third exposures to puns, humour was associated with a longer duration of time spent considering whether an item is still humorous after comprehension (referred to as rating duration). The association between longer rating durations and humour ratings on second and third exposure was taken as suggestive support for Wyer and Collins' (1992) hypothesized elaboration process, as a second function (beyond incongruity-resolution) that can produce humour. Given that puns with high latent semantic incongruity were significantly more humorous than puns with low latent semantic incongruity only for the first three exposures to a pun, it is hypothesized here that latent semantic incongruity can be viewed as an indicator of elaboration potential. Semantically incongruous puns should have a greater range of related concepts and associates that can come into play.

4.2 Current Operationalization of Elaboration

Participants were asked to list as many words as comfortably possible that come to mind when considering each of the two implied concepts contrasted in written puns. It was expected that the number of words that readily come to mind when thinking of the concepts from written puns (elaboration quantity) would be positively associated with both humour appreciation (mirth and cognitive appraisal), and also with latent semantic incongruity. The duration of time that participants were willing to spend on the elaboration task (elaboration duration) was also hypothesized to be positively associated with humour.

The comprehension-elaboration model claims that elaboration on the humorous aspects of stimuli should take place only after humorous incongruities have been adequately understood. However, prior studies have also demonstrated that priming content from humorous stimuli prior to exposure can also enhance humour (Goldstein, Suls, & Anthony, 1972; Topolinski, 2014). To test whether prior exposure is important to humour from elaboration, participants were assigned to complete the elaboration task either with the pun present (elaboration after viewing and comprehending each pun) or prior to viewing (and comprehending) each referent written pun. Participants should be able to provide a greater quantity of elaboration and find the puns to be significantly more humorous when the pun is present prior to completing the elaboration task (as opposed to elaboration without yet having been exposed to the referent pun).

4.3 Further Investigation of Semantic Incongruity

In the earlier studies conducted for this dissertation, semantic incongruity was investigated as a passive characteristic of puns (using Latent Semantic Analysis; Landauer, Foltz & Laham, 1998) and investigated in relation to key variables of interest. The extent to which participants actually engage with and recognize semantic incongruities should also be a meaningful predictor of humour appreciation. To test this hypothesis, in the current study, "semantic focus" was experimentally manipulated as an independent variable of interest. Participants were assigned to subjectively rate either the level of semantic similarity or the level of semantic dissimilarity between the two implied

concepts in each written pun. This manipulation was intended to focus participant attention on the semantic incongruity of the pun. If participant engagement with semantic incongruity is important to humour appreciation, then participants assigned to rate semantic dissimilarity should find the written puns to be more humorous than participants asked to rate semantic similarity. When treated as a dependent variable, subjective estimates of semantic dissimilarity should be correlated with latent semantic incongruity estimates and have a similar pattern of association with both humour ratings and with performance on the elaboration task.

4.4   Moderator Variables

Based on prior results from studies 1 and 2, latent semantic incongruity, aggression, and familiarity with a pun from a prior occasion were studied as potentially important moderators of the association between elaboration task variables and participant humour ratings according to both mirth (current emotional experience of humour in response to each item) and cognitive appraisal of humour (an overall objective assessment of the quality of humour in each item).

4.4.1   Aggression

Wyer and Collins (1992) argued that the association between aggressive content and humour appreciation should depend on its impact on elaboration. Aggression should either enhance or inhibit humour appreciation according to whether or not it facilitates cognitive elaboration. For example, you would be unlikely to engage and work through humorous stimuli if offended or upset by aggressive content. As such, the association between humour ratings and elaboration quantity and elaboration duration should be significantly moderated by the presence of aggressive content. If, on average, aggression facilitates engagement with pun items, then the positive association between elaboration and humour ratings should be stronger for puns with moderately aggressive content.

4.4.2   Familiarity

Prior empirical models have typically emphasized the importance of novelty and surprise in humour appreciation (Suls 1972; Schultz, 1972; Wyer & Collins, 1992; Ritchie, 2004).

The strong positive association between familiarity with a pun from a prior occasion and humour ratings in study 1 was therefore unexpected. The current study investigates the possibility that familiarity is positively associated with humour because it facilitates humour from elaboration.

4.4.3 Latent Semantic Incongruity

Evidence from prior studies supports the hypothesis that latent semantic incongruity is important to humour from incongruity comprehension (humour from comprehension of incongruities for puns with a low level of familiarity). In study 1, puns with high latent semantic incongruity were more humorous than puns with low latent semantic incongruity only for puns with a low average level of familiarity. In study 2, puns with high latent semantic incongruity were more humorous than puns with low latent semantic incongruity only for the first three exposures to a pun. The current study investigated whether semantic incongruity is important to humour from elaboration. Based on the results from study 2, it was hypothesized that puns with higher levels of latent semantic incongruity should have a greater quantity of associated novel related semantic features and content that can come into play.

4.4.4 Hypotheses

4.4.4.1 Elaboration and Humour

According to the comprehension-elaboration model (Wyer & Collins, 1992) both elaboration quantity (the number of associated words which come to mind in association with the two concepts in each pun) and elaboration duration (the total duration of time that the participants are willing to spend on the elaboration task for each pun) should be positively associated with humour ratings. In contrast, it could be speculated that the fluency account (Topolinski, 2014) would instead hold that a greater quantity of elaborations in a *shorter* elaboration duration (and therefore fluent elaboration) should be associated with humour ratings.

4.4.4.2   Pun Presence

The comprehension-elaboration model (Wyer & Collins, 1992) holds that elaboration should be positively associated with humour only when it takes place after comprehension of incongruity. Therefore, participants who were assigned to perform the elaboration task on concepts from each pun with the pun present should find the written puns to be significantly more humorous than participants who performed the elaboration task prior to viewing each pun. Participants who completed the elaboration task with the pun present should also have significantly greater elaboration quantity and they should be willing to spend longer durations of time on the elaboration task.

4.4.4.3   Semantic Focus

Studies 1 and 2 treated semantic dissimilarity as a passive latent characteristic of puns. However, it should be the extent to which participants actively attend to and engage with semantic incongruities that should be predictive of humour. To test this hypothesis, participants were asked to rate either the level of semantic dissimilarity or the level of semantic similarity between the implied concepts in each pun. Participants asked to rate semantic dissimilarity (thus emphasizing incongruity) should find the puns to be significantly more humorous than participants asked to rate semantic similarity.

When examined as a dependent variable, participant semantic relatedness ratings should be significantly positively correlated with latent semantic incongruity, humour ratings, and with both elaboration quantity and duration.

4.4.4.4   Elaboration and Moderator Variables

Wyer and Collins (1992) hypothesized that elaboration should be an important predictor of humour for stimuli that is familiar from a prior occasion. The positive relationship between familiarity with a pun from a prior occasion and humour (as identified in study 1) should therefore be dependent on elaboration.

Latent semantic incongruity was also investigated as a potential moderator of the relationship between elaboration and humour. Based on results from study 2, it was speculated that latent semantic incongruity may be an indicator of the potential of a pun

for the generation of novel elaborations. If so, then latent semantic incongruity should be positively associated with elaboration. Latent semantic incongruity was also investigated as a moderator of the association between elaboration and humour: investigating whether elaboration can explain the positive association between latent semantic incongruity and humour ratings (or if semantic incongruity can explain a relationship between elaboration and humour ratings).

Study 1 found that aggression functioned as a source of humour that was independent from latent semantic incongruity. Aggressive content should therefore be positively associated with humour and it may facilitate elaboration task performance. Aggression was also investigated as an exploratory moderator of the association between elaboration and humour. That is, aggressive content may facilitate humour from elaboration or elaboration may facilitate humour from aggressive content.

4.5 Method

4.5.1 Participants

In total, 199 participants were recruited at the University of Western Ontario, but 9 participants were removed from analyses as outliers (due to low item-total correlation with responses from other participants). Therefore, data from, 190 undergraduate students (65 females, 125 males; mean participant age = 18.23, SD = 1.79) were included in the reported analyses. Participants either spoke English as their first language (187 participants) or they had been speaking English as their primary language for at least 10 years (3 participants; mean years of English experience = 10.33; SD = .58). Students who participated in studies 1 or 2 were not eligible to participate in study 3.

There were two between-subject independent variables examined in the current study. Participants were assigned to rate all puns for the level of semantic similarity (47 participants completed item list 1; 48 completed list 2) or the level of semantic dissimilarity (50 participants completed list 1; 45 participants completed list 2) between the two concepts in each pun. Participants were also assigned to complete the elaboration task with either the pun present (44 participants completed list 1; 46 participants completed list 2) or to complete the elaboration task prior to having seen the referent pun

(53 participants completed list 1; 47 participants completed list 2). Although this is functionally a 2x2 experimental design, the interaction between these two variables are not reported because the theoretical rationale was based on targeted hypotheses. Accordingly, a priori *t*-tests were conducted to compare the two levels of each independent variable.[7]

4.5.2 Materials

The experiment was completed using the same custom html Internet survey platform developed by Dr. Rod Martin that was used in studies 1 and 2. Participants completed the experiment over the Internet on their own computers at a time and place of their choosing within 1 week of signing up for the experiment.

4.5.2.1 Items

To minimize demands on participant time and energy, participants were assigned to work through only one of two lists of 50 puns (the experiment took approximately 45 minutes to work through a list of 50 pun items). All pun items were originally collected from public submission databases available at punsandjokes.com and punoftheday.com. In order to minimize ethical concerns, no puns with racist, sexist, offensive, or excessively aggressive content were selected.

In total, one hundred pun items were selected: 34 based on homophone, 33 based on homograph, and 33 based on rhymes. Pun items were selected from those used in study 2 based on mirth and cognitive appraisal ratings (33% from those with low humour ratings, 34% from those with medium humour ratings, and 33% from those with high humour ratings). All study 3 items were selected from those used in study 2. The complete list of study 3 items is available in Appendix A.

---

[7] Nonetheless, for the sake of completeness, preliminary 2x2 analyses were conducted to examine whether these two variables interact. There was no significant interaction between semantic focus (and elaboration before or after viewing the referent pun according to effective humour factor loadings, elaboration quantity and elaboration duration.

4.5.3   Latent Semantic Incongruity

The same measure of latent semantic incongruity that was developed in study 1 and applied to study repetition effects in study 2 was employed once again in the current study. Example dictionary definition passage vectors and LSA estimates are available in Appendix B.

4.5.4   Procedure

The study proceeded as follows. Prior to starting the experiment participants provided demographics information. For each pun, participants first completed the elaboration task (with the actual referent pun either absent or present), in a subsequent screen they provided semantic relatedness ratings (either the level of semantic dissimilarity or semantic similarity between the implied concepts in each pun), and in a final screen they rated each pun according to their experience of mirth, their cognitive appraisal of the humour in the pun, the level of aggressive content, and familiarity with the pun from a prior occasion.

4.5.4.1   Demographics

Participants were first asked to provide general demographics: age, gender, whether English was their first language and (if English was not their first language) the number of years they have been speaking English as their primary language.

4.5.4.2   Elaboration Task

Participants received on-screen instructions to list as many words as they were comfortably able to provide that came to mind when considering each of the two implied concepts in each pun. Two example trials were presented, which displayed a screenshot of a completed elaboration with four associated words provided for each of the concepts (see Appendix C for the example trials as provided to participants). Next, participants were presented with one of the two implied concepts from a pun: concept A (the first implied concept as introduced in the pun sentence) or concept B (the alternate concept implied by the key ambiguous word in the sentence. The order in which concept A and

concept B were presented for elaboration were randomized for each item between participants.

For each pun item, a single word describing concept A (or concept B; randomized as to which was presented first) with one to two disambiguating synonyms would appear above a text box. Participants were asked to provide as many words that come to mind (as many as they are comfortable or willing to provide) in association with the presented concept in the text box. For example, they might have been shown the concept "syncing: electronics" (concept A). Once they were satisfied with their elaboration word list for this concept, they could press the "next" button to bring up a second implied concept cue. To continue the example, they would have been presented with the concept "sinking: submerge" (concept B). Once satisfied, they could press the "next" button again to move onto the (dis)similarity rating task. These example concept cues are taken from the pun "I changed my Ipod's name to the Titanic and now its syncing".

Two variables of interest were prepared from the elaboration task. The average number of words provided during the elaboration task for each pun (the average word count across both concepts A and B) was recorded as "elaboration quantity". The average quantity of time spent on the elaboration task (also across both concepts A and B) for each pun was recorded as "elaboration duration".

4.5.4.3   Semantic Focus

After completing the elaboration task participants were presented with the two implied concepts from each pun (e.g. syncing, sinking). Participants were each randomly assigned to rate all items for either the level of semantic similarity (on a scale from 0 not at all similar to 6 very similar) or the level of semantic dissimilarity between the two concepts in each pun (on a scale from 0 not at all dissimilar to 6 very dissimilar). The on-screen instructions asked participants to rate the level of semantic dissimilarity (or the semantic similarity) between the two presented concepts.

4.5.4.4   Pun Presence

Participants were assigned to work through the entire experiment either with the pun present during both semantic relatedness rating and the elaboration task (therefore, performing the elaboration task after viewing and comprehending the pun), or to a condition in which the pun was not presented until after both semantic relatedness rating and the elaboration task (therefore, performing the elaboration task prior to viewing and comprehending the referent pun).

4.5.4.5   Pun Ratings

After providing ratings of the semantic relatedness between the two concepts in each pun, participants were presented with four Likert-type scales. They were asked to rate the extent to which each pun caused them to experience the emotional state of mirth (from 0 not at all to 6 strong experience of mirth), their cognitive appraisal of each pun (from 0 not at all to 6 strong cognitive appraisal), the extent to which each pun contains aggressive themes (from 0 not at all to 6 very aggressive), and the extent to which they are familiar with each item from a prior occasion (on a scale from 0 not at all to 6 very familiar). These four Likert-type rating scales appeared in a random order between participants. A detailed definition of mirth and cognitive appraisal of humour was available on the bottom of the rating screen for all puns. Mirth was defined as "the *emotional* experience of amusement, fun, up to hilarity; typically accompanied by the urge to grin, smile or laugh" and cognitive appraisal was defined as "being clear, clever, and an overall good example of humour; something that you think other people would enjoy". Participants were allowed to select "no answer" for any of these scales if they did not feel comfortable rating a pun. After providing these ratings, they could press the "next" button (which would bring up the elaboration task for the next item.

4.5.5   Data Preparation

4.5.5.1   Average Over Items

Inferential statistics were conducted on data that was averaged across items. Item averages were prepared for mirth, cognitive appraisal, aggressiveness, familiarity ratings,

and semantic relatedness ratings. Averages over items were also prepared for the overall average elaboration quantity and elaboration duration for each item.

4.5.5.2   Missing Data

Elaboration quantity data was accepted and unchanged regardless of the number of words provided (even if zero associated words were provided). Elaboration duration data was flagged as missing and replaced with the average latency across participants for trials with elaboration task duration greater than 8 minutes (480000ms; approximate cut off chosen based on 97.5th percentile). Missing data due to this criterion for excessively long duration had to be replaced only rarely (366; 3.9% of 9500 cases). Missing data due to participant selection of the "no answer option" was also replaced with the mean across participants for that item for mirth (35 cases; 0.4% of 9500 cases), cognitive appraisal (32 cases; 0.3% of 9500 cases), familiarity (29 cases; 0.3% of 9500 cases), aggression (50 cases; 0.5% of 45 cases), and similarity ratings (45 cases; 0.5% of 9500 cases).

4.5.5.3   Internal Reliability

Participant item-total correlations were calculated using Cronbach's alpha (with participant responses treated as items on a scale). Participants with low item total correlations according to any one scale (less than .2) were removed from analysis as atypical outliers. Based on this cut-off criterion, 9 people were removed from further analyses. Thus, although 199 people were recruited to participate the current study in total, the analyses reported below are based on the remaining 190 participants (97 participants were assigned to work through list 1 containing 50 puns; 93 participants worked through list 2 which contained a second set of 50 puns).

With 9 participants removed, participant reliability was high for ratings of mirth (list 1 α = .95, list 2 α = .95), cognitive appraisal of humour (list 1 α = .95, list 2 α = .95), for the presence of aggressive themes (list 1 α = .97, list 2 α = .97), familiarity from a prior occasion (list 1 α = .96, list 2 α = .97) and for ratings of both subjective similarity (list 1 α = .93, list 2 α = .96) and dissimilarity (list 1 α = .97, list 2 α = .96) between the concepts in each pun. There was also high reliability, over participants, for elaboration quantity; the number of words provided per elaboration in response to concepts A and B (list 1 α =

.82, list 2 α = .81), but participant reliability was relatively low for the duration of time participants spent on the elaboration task for concept A and B (list 1 α = .35, list 2 α = .65). For the complete list of reliability estimates, refer also to Table 11.

4.6 Results

4.6.1 Dimension Reduction

Mirth and cognitive appraisal were very highly correlated $r(98) = .99$, $p < .001$ and there were no meaningful differences in the pattern of results between mirth and cognitive appraisal ratings of humour. As such, mirth and cognitive appraisal ratings were collapsed into a single outcome variable of interest using a principle component factor analysis. The mean scores (over items) for mirth and cognitive appraisal were entered into the factor analysis. A single factor solution could account for 99.6% of the variability in the data (as in study 1), and the factor was again named the "effective humour" factor. The extent to which each pun loads on the effective humour factor was saved using the regression method to be used as an outcome variable of interest in subsequent analyses.

4.6.2 Elaboration and Humour

Overall (regardless of whether or not the pun has been viewed yet), fluent elaboration was positively associated with humour. That is, the quantity of words provided in the elaboration task was positively associated with effective humour $r(98) = .28$, $p < .01$ and the duration of time spent on the elaboration task was negatively associated with effective humour $r(98) = -.24$, $p < .05$. Effective humour was positively associated with both familiarity $r(98) = .56$, $p < .001$ and with aggression $r(98) = .39$, $p < .001$ (replicating findings from study 1). For the complete set of bivariate correlations, see Table 12.

Table 11: Cronbach's alpha reliability estimates for participants for each study 3 dependent variable broken down by two item lists.

	List 1		List 2	
	N	α	N	α
Cognitive Appraisal	97	0.95	93	0.95
Mirth	97	0.95	93	0.95
Familiarity	97	0.96	93	0.97
Aggression	97	0.97	93	0.97
Dissimilarity Rating	50	0.97	45	0.96
Similarity Rating	47	0.93	48	0.96
Elaboration Quantity	97	0.82	93	0.81
Elaboration Duration	97	0.35	93	0.65

Note: Participant sample size (N) for each dependent variable is reported.

Table 12: Bivariate correlations between dependent variables for the study 3 pun items.

	1	2	3	4	5	6	7	8	9	10
1. Latent Semantic Incongruity	-	.30**	.26**	.28**	-0.02	.28**	.22*	-0.17	.18	-.14
2. Effective Humour Factor		-	.996**	.996**	.39**	.56**	.28**	-0.24*	-.04	.14
3. Cognitive Appraisal			-	.986**	.39**	.57**	.29**	-.24*	.006	.09
4. Mirth				-	.39**	.55**	.27**	-.23*	.03	.08
5. Aggression					-	0.08	0.19	-0.16	.16	-.10
6. Familiarity						-	0.01	-0.16	-.98	.16
7. Elaboration Quantity							-	-.55**	-.08	.16
8. Elaboration Duration								-	-.35**	.27**
9. Dissimilarity Ratings									-	-.94**
10. Similarity Ratings										-

Note: Correlations marked by (*) were statistically significant at $p < .05$, correlations marked by (**) were statistically significant at $p < .01$.

4.6.2.1 Pun Presence

Elaboration prior to viewing the referent pun was compared with elaboration with the pun present accorded to the dependent variables (collapsed across semantic focus conditions). Participants who were asked to complete the elaboration task with the pun present found the puns to be significantly more humorous $t(99) = 7.85$, $p < .001$, than participants who completed the elaboration task prior to viewing the referent written pun. Participants who completed the elaboration task with the pun present also provided significantly more words in the elaboration task $t(99) = 8.97$, $p < .001$ and completed the elaboration task in a significantly shorter duration of time $t(99) = 3.33$, $p < .01$. Averages over items broken down by pun presence condition are presented in Table 13.

4.6.3 Familiarity and Elaboration

A stepwise multiple regression was employed to investigate familiarity as a moderator of the association between elaboration quantity, elaboration duration and effective humour factor loadings. At the first step, there was no significant main effect of elaboration quantity $\beta = .72$, *ns*, or for elaboration duration $\beta = -0.000014$, *ns*; $R^2 = .09$, $F(2, 97) = 4.83$, $p < .05$. At a second step there was a significant main effect of familiarity $\beta = -1.40$, $p < .001$ and elaboration quantity $\beta = .93$, $p < .01$, but not for elaboration duration $\beta = 6.5861E-7$, *ns*; $R^2 = .40$, F change $(1, 96) = 48.34$, $p < .001$. At the final step, there was a significant main effect of familiarity $\beta = -14.69$, $p < .05$ and elaboration duration $\beta = -0.000062$, $p < .05$, but not for elaboration quantity $\beta = -.30$, *ns*. There were also significant moderating interactions of familiarity on the association between effective humour and both elaboration quantity $\beta = 1.32$, $p < .05$ and elaboration duration $\beta = 0.000073$, $p < .01$; $R^2 = .45$, F change $(2, 94) = 4.81$, $p < .01$. These interactions are described in greater detail in the following subsections.

Table 13: Mean and standard deviation for study 3 puns by pun presence condition.

Dependent Variable	Pun Absent Mean (*SD*)	Pun Present Mean (*SD*)	Overall Mean (*SD*)
Cognitive appraisal	2.40 (0.63)	2.71 (0.58)	2.56 (.58)
Mirth	2.28 (0.65)	2.47 (0.60)	2.37 (.60)
Familiarity	0.79 (0.40)	0.88 (0.43)	0.83 (.40)
Aggression	1.15 (0.60)	1.41 (0.56)	1.28(.56)
Dissimilarity Rating	4.98 (0.62)	5.10 (0.65)	5.05 (.61)
Similarity Rating	.78 (0.65)	.98 (0.62)	0.88 (.62)
Elaboration Quantity	6.78 (0.32)	7.07 (0.36)	6.93 (.30)
Elaboration Duration	97353.22 (10096.79)	101899.40 (11327.60)	99626.31 (8278.69)
Latent Semantic Incongruity	.49 (.13)	.49 (.13)	.49 (.13)

Note: Mean and standard deviation (SD) over 100 puns are reported broken down by pun presence condition: elaboration with the pun present (90 participants), or elaboration prior to viewing the referent pun (pun absent; 100 participants).

4.6.3.1 Familiarity and Elaboration Quantity

Post hoc simple tests of slopes were conducted to describe the moderating effect of familiarity on the association between elaboration quantity and humour. Puns with a higher quantity of elaboration (set as 8 words; mean elaboration quantity = 6.95, SD = .30) had higher effective humour factor loadings than puns with a lower quantity of elaboration (set as 6) only for puns with a moderate level of familiarity from a prior occasion $t(99) = 2.90$, $p < .01$ (set as 2). For puns with a lower level of familiarity from a prior occasion (set as 0.5), there was no significant effect of elaboration quantity on effective humour.[8] In summary, there was an association between elaboration quantity and humour only for puns that were moderately familiar from a prior occasion. The interaction between familiarity and elaboration quantity on effective humour factor loadings is illustrated in Figure 6.

4.6.3.2 Familiarity and Elaboration Duration

Further post hoc simple tests of slopes revealed that effective humour factor loadings were higher for items with longer elaboration duration (set as 116183 ms; approximately 2 standard deviations above the mean) than for puns with shorter elaboration duration (Set as 83068 ms; approximately 2 SD below the mean) only for puns with a moderate level of familiarity from a prior occasion, $t(99) = 2.59$, $p < .05$ (set again as 2). For puns with a lower level of familiarity from a prior occasion (set again as 0.5), there was no significant effect of elaboration duration; albeit there was a trend towards lower elaboration duration being more humorous than longer elaboration durations, $t(99) = -1.77$, $p < .08$. In summary, there was an association between elaboration duration and humour only for puns that were moderately familiar from a prior occasion. The interaction between familiarity and elaboration duration for effective humour factor loadings is illustrated in Figure 7.

---

[8] Average familiarity ratings for study 3 puns were relatively low and only moderately familiar at most; ranging from 0.21 to 2.51; $M = 0.83$, $SD = .40$
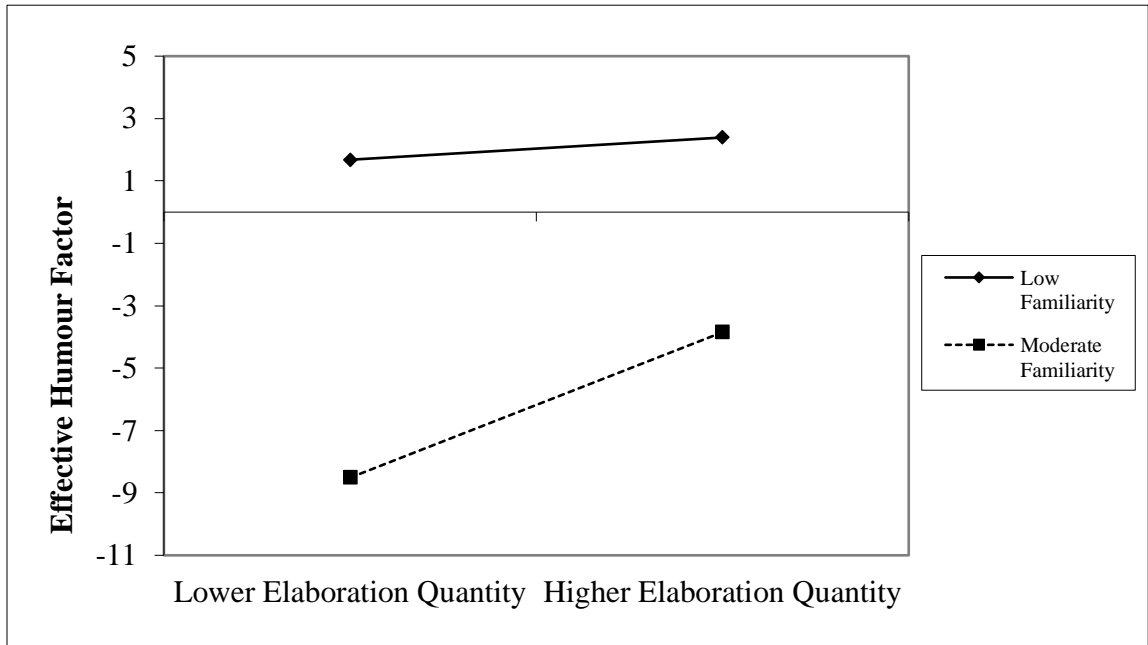
Figure 6: Moderating effect of familiarity on the relationship between elaboration quantity and the effective humour factor in study 3.

Puns with a higher level of elaboration quantity (set as 8 words) were more humorous than puns with a lower level of elaboration quantity (set as 6 words) only for puns with a moderate level of familiarity (low familiarity set at 0.5, moderate familiarity set at 2).
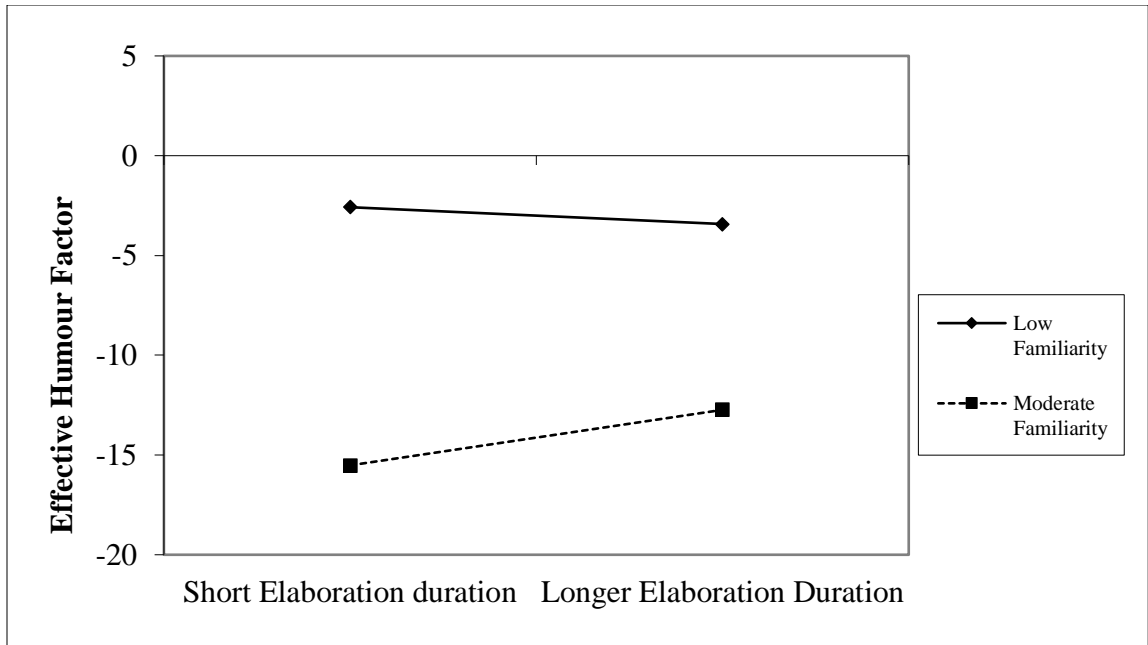
Figure 7: Moderating effect of familiarity on the relationship between elaboration duration and the effective humour factor in study 3.

Puns with higher elaboration task duration (set as 116183.69 ms) were more humorous than puns with lower elaboration duration (set as 83068.93 ms) only for puns with a moderate level of familiarity (low familiarity set at 0.5, moderate familiarity set at 2.0).

4.6.4   Aggression and Elaboration

Aggression was also investigated as a potential moderator of the association between elaboration quantity, elaboration duration and effective humour factor loadings using a stepwise regression approach. At the first step, there was no significant main effect of elaboration quantity $\beta = .72$, *ns,* or of elaboration duration $\beta = -0.000014$ *ns*; $R^2 = .09$, $F(2, 97) = 4.83$, $p < .05$. At the second step, there was a significant main effect of aggression $\beta = .61$, $p < .001$, but there was no significant main effect of elaboration quantity $\beta = .56$, *ns,* or elaboration duration $\beta = -0.000011$, *ns*; $R^2 = .20$, F change $(1, 96) = 13.52$, $p < .001$. At the final step, there was a significant main effect of elaboration duration $\beta = -0.000077$, $p < .05$, but there was no significant main effect of elaboration quantity $\beta = .61$, *ns,* or aggression $\beta = -3.15$, *ns*. At the final step there was also a significant moderating interaction of aggression on the association between elaboration duration and humour $\beta = 0.00005$, $p < .05$, but there was no significant interaction between aggression and elaboration quantity $\beta = -.15$, *ns*; $R^2 = .27$, F change $(2, 94) = 4.49$, $p < .05$.

Post hoc simple tests of slopes revealed that shorter elaboration durations (again set at 83068 ms) were more humorous than longer elaboration durations (set at 116183 ms) only for puns with a low level of aggressive content (set as 0.5; on a scale of 0-6), $t(99) = -2.37$, $p < .05$. For puns with a moderate level of aggressive content (set as 2)[9], longer elaboration durations trended towards being more humorous than shorter elaboration durations $t(99) = 1.89$, $p < .06$. In summary, longer elaboration is beneficial to effective humour for puns with aggressive content, whereas shorter elaboration is beneficial to effective humour for puns without clear aggressive content. The interaction between aggression and elaboration duration for effective humour factor loadings is illustrated in Figure 8.

---

[9] Average aggression levels were also relatively low $M = 1.28$ $SD = .56$
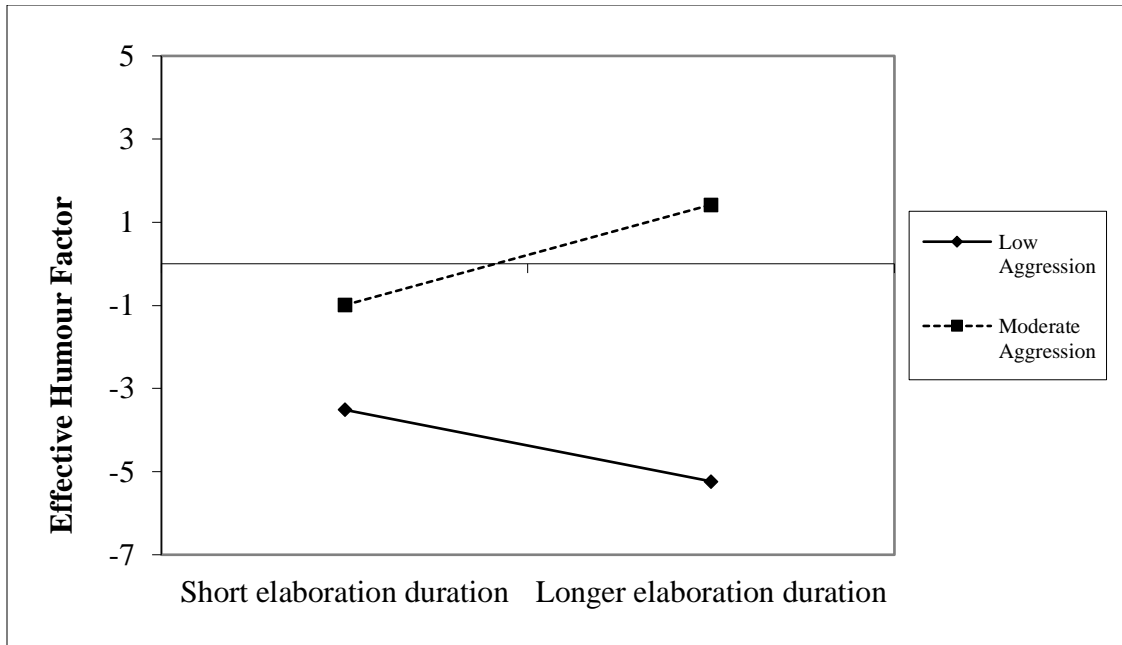
Figure 8: Moderating effect of aggression on the relationship between elaboration duration and loadings on the effective humour factor in study 3.

Puns with high elaboration duration (set as 116183.69 ms; 2 SD above the mean) were more humorous than puns with lower elaboration duration (set as 83068.93 ms; 2 SD below the mean) only for puns with a moderate level of aggression (low aggression set at 0.5, moderate aggression set at 2.0; on a likert type scale from 0-6). For puns with low aggression, low elaboration duration was trending towards being more humorous than puns with high elaboration duration ($p < .06$)

4.6.5   Semantic Incongruity and Elaboration

4.6.5.1   Latent Semantic Incongruity

Latent semantic incongruity was significantly positively correlated with effective humour factor loadings $r(98) = .30$, $p < .01$, elaboration quantity $r(98) = .22$, $p < .05$, and familiarity $r(98) = .28$, $p < .05$, but not with elaboration duration $r(98) = -.17$, *ns,* or aggression $r(98) = -.02$, *ns*; refer to Table 12.

Latent semantic incongruity was investigated as a potential moderator of the association between elaboration quantity, elaboration duration and humour. At the first step, there was no significant main effect of elaboration quantity $\beta = .72$, *ns,* or of elaboration duration $\beta = -0.000014$, *ns*; $R^2 = .09$, $F(2, 97) = 4.83$, $p < .05$. At the second step, there was a significant main effect of latent semantic incongruity predicting humour $\beta = 1.66$, $p < .05$, but there was no significant effect of elaboration quantity $\beta = .58$, *ns,* or elaboration duration $\beta = -1.294\text{E-}5$, *ns*; $R^2 = .13$, $F(1, 96) = 4.57$, $p < .05$. At the final step, there were no statistically significant moderating interactions and interaction terms did not significantly improve model fit, $R^2 = .15$, F change $(2, 94) = 1.11$, *ns*. Therefore, the effect of latent semantic incongruity on effective humour factor loadings was not significantly moderated by elaboration task variables.

4.6.5.2   Semantic Focus

Participant focus on semantic dissimilarity (as opposed to semantic similarity) was examined as an independent variable of interest (collapsed across pun presence elaboration conditions). Paired sample *t*-tests (accounting for variability between items) were conducted to compare dependent variables of interest for participants assigned to rate semantic dissimilarity with participants assigned to rate semantic similarity. Participants rating semantic dissimilarity found the puns to be significantly more humorous $t(99) = 5.19$, $p < .001$. In contrast, elaboration quantity was significantly higher for participants asked to rate semantic similarity $t(99) = -10.37$, $p < .001$. Mean scores averaged over items broken down by semantic focus condition is available in Table 14.

Table 14: Study 3 mean and standard deviation over items broken down by semantic focus condition.

Dependent Variable	Similarity Focus Mean (*SD*)	Dissimilarity Focus Mean (*SD*)	Overall Mean (*SD*)
Cognitive Appraisal	2.45 (0.63)	2.66 (0.58)	2.56 (.58)
Mirth	2.27 (0.62)	2.48 (0.61)	2.37 (.60)
Familiarity	0.78 (0.40)	0.88 (0.46)	0.83 (.40)
Aggression	1.21 (0.60)	1.35 (0.57)	1.28(.56)
Dissimilarity Rating	-	5.04 (0.61)	5.05 (.61)
Similarity Rating	0.88 (.61)	-	0.88 (.62)
Elaboration Quantity	7.06 (0.34)	6.7866 (0.32)	6.93 (.30)
Elaboration Duration	99866.56 (10432.22)	99386.05 (10413.99)	99626.31 (8278.69)
Latent Semantic Incongruity	.49 (.13)	.49 (.13)	.49 (.13)

Note: Study 3 mean and standard deviation (*SD*) over 100 items. Participants were asked to either rate either the level of semantic dissimilarity (95 participants) between the two concepts each pun or the level of semantic similarity (95 participants).

4.6.5.3   Semantic Relatedness Ratings

There was a ceiling effect for subjective semantic dissimilarity estimates (overall mean dissimilarity = 5.08, SD = .61; on a Likert-type scale from 0 to 6) and a floor effect for semantic similarity estimates (overall mean similarity = .89, SD = .61; see Table 14). On average, the two concepts in each pun were subjectively viewed as being nearly completely dissimilar to each other. Participant Likert-type scale semantic similarity ratings were highly correlated with semantic dissimilarity ratings $r(98)$ = -.94, $p < .001$. Participant estimates of semantic dissimilarity and semantic similarity were not significantly correlated with the effective humour factor, aggression, familiarity, elaboration quantity or latent semantic incongruity. Semantic dissimilarity ratings were associated with shorter elaboration task durations, $r(98)$ = -.31, $p < .01$ and semantic similarity estimates were associated with longer elaboration durations $r(98)$ = -.31, $p < .001$.

4.7   Discussion

A free association concept cueing task was employed to operationalize and test whether humour can be produced from two distinct cognitive processes: both comprehension of incongruities and from post-comprehension elaboration on the humorous aspects of stimuli. The primary goal of the current study was to test the comprehension-elaboration based prediction that elaboration can account for humour from familiar stimuli (Wyer & Collins, 1992). Semantic incongruity, aggression, and familiarity were investigated as moderators of the association between elaboration task performance and humour appreciation. Overall, humour ratings were positively associated with elaboration quantity, but also with shorter elaboration durations (indicating an overall positive function of fluency). However, consistent with predictions from comprehension-elaboration, moderating interactions revealed that elaboration quantity and *longer* durations of time spent on the elaboration task were positively associated with humour ratings, but only for puns with a moderate level of familiarity and for puns with a moderate level of aggressive content. There were no significant effects of elaboration for puns with a low level of familiarity from a prior occasion.

The second important goal of this study was to investigate the function of semantic incongruity in humour appreciation from elaboration. Consistent with the elaboration potential hypothesis, latent semantic incongruity was positively correlated with elaboration quantity. Whereas prior studies only investigated semantic incongruity as a passive latent characteristic of puns, the current study also experimentally manipulated participant attention to either semantic dissimilarity or semantic similarity in puns. Participants attending to semantic dissimilarities found the written puns to be more humorous than participants attending to similarities. However, it was unexpectedly also found that participants attending to semantic similarities produced a greater quantity of elaboration.

Recall that mirth (the emotional experience of humour) and cognitive appraisal (a more objective overall assessment of the quality of an item) could be meaningfully distinguished when written puns were rated multiple times in study 2. In that study, latent semantic incongruity significantly moderated decreases in mirth, but not cognitive appraisal with repetition. In the current study, participant assessment of their own emotional mirth response to the puns was nearly completely equivalent with the more objective "cognitive appraisal" of humour in written puns.[10] As such, the two humour assessment variables were combined using a principal components factor analysis and assessed as a single "effective humour" latent factor. The finding that there is no meaningful difference between mirth and cognitive appraisal in study 3 provides convergent support for the reliability of study 1 findings (in which participants were asked to assess the puns and not their current emotional response; several of these findings are successfully replicated in the current study).

4.7.1 Familiarity and Humour from Elaboration

Familiarity was again found to be strongly positively associated with effective humour factor loadings (replicating findings from study 1). The current results further suggest

---

[10] In the current study correlation strength between mirth and cognitive appraisal was high $r(98) = .996$.

that familiarity might be positively associated with humour ***because*** of post-comprehension elaboration. This conclusion is based on the finding that puns with high elaboration quantity were significantly more humorous than puns with low elaboration quantity, but only for puns with a moderate level of familiarity. The comprehension-elaboration model also predicted that longer durations of time participants were willing to spent on elaboration should be predictive of humour. Overall, larger elaboration quantities within shorter elaboration durations were positively associated with humour. However, more clearly consistent with this prediction, longer elaboration durations were associated with humour for puns with a moderate level of familiarity. There was no effect of elaboration (elaboration quantity or elaboration duration) for puns with a low level of familiarity from a prior occasion. Note that current results are limited by the comparatively restricted range in participant familiarity ratings (in comparison to study 1).[11]

4.7.2   Aggression and Humour from Elaboration

Overall, aggression was positively correlated with humour and trending towards a significant association with elaboration quantity (p < .06). Overall, shorter elaboration durations were positively associated with humour, but longer elaboration durations were positively associated with humour for puns with a moderate level of aggressive content (trending towards an association between shorter elaboration durations and humour for puns with low aggressive content). Participants may be more willing to spend a longer duration of time on the elaboration task for puns with aggressive content. Wyer and Collins (1992) predicted that aggression should be positively associated with humour only if it facilitates elaboration. Current results indicate that aggressive content is associated with humour from both comprehension challenge and from elaboration.

---

[11] In study 2, the dissociation between mirth and cognitive appraisal was only observed over repeated exposures; whereas, study 3 average familiarity ratings for the written pun items were relatively low mean familiarity = 0.83 *SD* = .40 (on a scale from 0-6).

4.7.3   Elaboration After Comprehension

Participants who elaborated on associated concepts after viewing (and comprehending) the referent pun (as opposed to prior to viewing the pun) provided a greater quantity of elaborations in a shorter duration of time and found the puns to be more humorous. Wyer and Collins (1992) argued that humorous elaboration should take place only after humorous incongruities have been adequately understood. Elaboration was hypothesized to be associated with humour more strongly when performed after viewing and comprehending the referent stimuli. However, prior studies have also found that priming content from humorous stimuli prior to exposure can also enhance humour (Goldstein, Suls, & Anthony, 1972; Topolinski, 2014). In support of Wyer and Collin's hypothesis, participants who performed the elaboration task after viewing the written puns found them to be significantly more humorous than participants who completed the elaboration task prior to comprehension. Elaboration after comprehension was also more productive: participants who performed the elaboration task after comprehension had significantly greater elaboration quantity and completed the task with shorter elaboration durations. These results are also consistent with the aforementioned finding that elaboration is associated with humour only for familiar stimuli (which therefore would also have been viewed on a prior occasion).

4.7.4   Semantic Focus

It was argued here that it should be the extent to which participants are willing and able to engage with the potential semantic incongruities of an item that should be predictive of humour. Consistent with this hypothesis, participants assigned to subjectively rate all items for semantic dissimilarity found the puns to be significantly more humorous than participants assigned to rate all items for semantic similarity. When examined as a dependent variable of interest there was a ceiling effect, such that participants found the two implied concepts in written puns to be nearly completely dissimilar to each other. The concepts were rated as being nearly completely semantically dissimilar to each other regardless of whether or not they had actually seen the pun (refer to Table 13; that is, there was no significant difference in semantic dissimilarity ratings based on pun presence condition). Latent semantic incongruity was not significantly correlated with

participant semantic relatedness ratings. Further research investigating the external validity of latent semantic incongruity as a measure of semantic dissimilarity may therefore be warranted.

There was an unexpected and potentially important advantage for participants assigned to rate semantic similarity. Participants assigned to rate all items for semantic similarity provided a greater quantity of words during the elaboration task. In their study of semantic relatedness as a predictor of participant humour ratings for metaphor stimuli, Hillson and Martin (1994) obtained an interaction effect that was analogous to results from the current study. Their "domains-interaction" approach to measuring semantic relatedness allowed them to assess both semantic dissimilarity and semantic similarity between the concepts in metaphors. They found that the most humorous metaphors were high in *both* semantic similarity and dissimilarity. Given current findings, semantic dissimilarity may produce humour from incongruity-resolution (humour from incongruity for puns with low familiarity) while semantic similarity may have enhanced or produced humour from elaboration (given that similarity focus was associated with a greater quantity of elaboration).

Latent semantic incongruity may have actually done a better job at estimating incongruity effects in humour than did participant estimates of semantic dissimilarity. Manipulating participant attention to semantic dissimilarity significantly increased participant humour ratings and focusing attention on semantic similarity improved performance on the elaboration task. However, there was a problematic ceiling effect for participant subjective estimates of semantic dissimilarity. On average, the participants found the two concepts in the puns to be nearly completely dissimilar to each other. Latent semantic incongruity, as a computational approach, is not subject to this potential participant bias.

Latent semantic incongruity was positively associated with elaboration quantity, but latent semantic incongruity did not significantly interact with elaboration quantity or elaboration duration in predicting humour. That is, latent semantic incongruity was positively associated with effective humour factor loadings regardless of performance on the elaboration task. Thus, the hypothesis that latent semantic incongruity was associated

with humour on repetition because it facilitated elaboration was not supported by the current results. Recall that in study 1 latent semantic incongruity was positively associated with humour only for items with a low level of familiarity from a prior occasion. In the current study, it was found that elaboration is associated with humour only for items with a moderate level of familiarity. Therefore, semantic incongruity appears to produce humour from stimuli with a low level of familiarity (presumably through an incongruity-resolution or comprehension process) whereas elaboration appears to produce humour from puns with a moderate level of familiarity from a prior occasion.

Future studies investigating the effect of semantic dissimilarity and similarity in humour could address the ceiling effect for semantic dissimilarity estimates by guiding them through an elaborative process. Wyer and Collins (1992) argued that humorous elaboration does not necessarily have to be internally generated. For example, comedians or humorous movies typically guide audiences through the humorous implications of events. Experimentally manipulating both elaboration quantity and participant attention to semantic (dis)similarities should therefore be possible by presenting participants with a prepared list of semantic similarities or dissimilarities between the concepts in puns.

4.7.5   Study 3 Summary

There was support for several core predictions from the comprehension-elaboration model, but with several important exceptions. Consistent with comprehension-elaboration hypotheses, elaboration quantity and longer elaboration durations were significantly associated with humour, but only for puns with a moderate level of familiarity from a prior occasion or for puns with aggressive content. Humour appreciation for novel puns and for moderately familiar puns appears to be produced by two processes: either from (a) comprehension fluency and semantic incongruity for low familiarity items (as in study 1) or (b) from elaboration for puns with a moderate level of familiarity. There was also support for the comprehension-elaboration based hypothesis that elaboration after viewing and comprehending an item is important for humour. Humour ratings and elaboration task performance were higher when the task was completed after viewing the pun (as opposed to prior to viewing and comprehending the referent pun).

Latent semantic incongruity was significantly positively correlated with both humour ratings and elaboration quantity, suggesting that it can potentially serve as a meaningful indicator of elaboration potential. However, semantic incongruity did not significantly moderate the association between elaboration quantity or duration and humour. Therefore, elaboration was not associated with humour *because* of semantic incongruity. Given results from studies 1-2, latent semantic incongruity appears to create humour primarily through comprehension fluency. The final important finding was that both semantic dissimilarity and semantic similarity appear to have a meaningful role in humour appreciation. Participants asked to rate semantic dissimilarity for the concepts from the puns found them to be significantly more humorous, while participants asked to rate semantic similarity produced a significantly greater quantity of elaboration (which in turn was significantly associated with humour for moderately familiar puns (see also Hillson & Martin, 1994).

Chapter 5

5   General Discussion

Explanatory models of humour appreciation have frequently emphasized the importance of incongruity. However, incongruity itself has been challenging to operationalize for experimental study (for an in depth discussion see Ritchie, 1999, 2004; 2009; Forabosco, 1992). The studies conducted for this dissertation operationalized incongruity for empirical study as the latent semantic dissimilarity between the two alternate implied concepts in puns using latent semantic analysis (Landauer, Foltz & Laham, 1998). Study 1 developed the latent semantic incongruity measure and compared puns based on homophones, homographs and rhymes. Aggression and familiarity were examined as moderators of the association between latent semantic incongruity and participant humour ratings. Study 2 investigated humour appreciation on repeated exposures to puns as a function of latent semantic incongruity, the duration of time necessary to "get" the humour in each pun (comprehension duration), the duration of time spent considering humour ratings in repeated pun items (rating duration), and according to how humour was assessed; according to either mirth (current emotional response to an item) or cognitive appraisal of humour a more objective overall appraisal of pun quality). Study 3 examined elaboration as a predictor of humour appreciation in puns. Participant estimates of aggressive content and familiarity were once again examined as important moderator variables. Study 3 also subtly manipulated participant attention to semantic dissimilarities between the two implied concepts in puns as an independent variable of interest. Across these three studies, specific predictions from the incongruity-resolution model (Suls, 1972), the comprehension-elaboration model (Wyer & Collins, 1992), and the fluency account of humour appreciation (Topolinski, 2014) were evaluated. In this section, the implications of current findings for these models given current results from studies 1 to 3 will be considered in greater detail.

The most important finding from the currently reported studies was that the variables involved in humour appreciation depend on whether or not a pun is familiar from a prior occasion. Incongruity-resolution and fluency of incongruity-comprehension was important to humour appreciation for low familiarity stimuli or for stimuli at first

exposure (studies 1-2), whereas elaboration quantity and elaboration duration was important to humour appreciation for familiar stimuli (study 3; where there was an effect of elaboration only for familiar items). Recall also that in study 2, fluency of first exposure comprehension was positively associated with humour ratings on repeated exposures. This finding suggests that people must have been either able to recall their comprehension fluency from the pun at first exposure (consistent with findings from Suls, 1975) or deduce how challenging each item must have been. This finding arguably supports the comprehension-elaboration hypothesis that humour appreciation from these two processes (comprehension fluency and elaboration) are not necessarily mutually exclusive (Wyer and Collins, 1992): elaboration was predictive of humour for familiar stimuli, but comprehension at first exposure also appears to continue to play a role in humour appreciation on repetition.

The most important exception to predictions from previous models of humour was regarding the importance of fluency in humour appreciation. Incongruity-resolution (Suls, 1972) and comprehension-elaboration (Wyer & Collins, 1992) emphasized that a moderate level of challenge (neither too easy nor too difficult to understand) should be optimal for humour appreciation. However, across all three studies there was no evidence of a non-linear (inverted-U shaped) relation between processing duration variables and humour. Written puns are considered to be relatively easy examples of humour (Wyer & Collins, 1992); therefore, greater complexity of problem-solving, effort, and longer durations of time necessary to adequately understand humorous stimuli could also have been predictive of humour. However, consistent with the fluency account (Topolinski, 2014; see also Goldstein, 1970a), both rapid and accurate identification of an item as a pun (study 1), and shorter durations of time necessary to "get" the humour in each item (study 2) was associated with humour. Study 3 found that fluent elaboration (overall; greater elaboration quantity in shorter elaboration durations) was positively associated with humour appreciation; with the exception of puns with a moderate level of aggressive content or for puns that are moderately familiar from a prior occasion. Going forward, explanatory accounts of humour appreciation must be able to accommodate a role for fluency in humour appreciation. Rather than comprehension difficulty, the most concise

explanation for current results must be that fluent comprehension of incongruities (easy, accurate, and rapid) from unfamiliar stimuli produces humour.

5.1.1   Latent Semantic Incongruity

Across three studies, latent semantic incongruity proved to be an important predictor of humour ratings. Several novel findings regarding the function of incongruity and humour were obtained using this new measure. Latent semantic incongruity was associated with an overall "effective humour" factor in study 1, but only for puns with a low level of familiarity from a prior occasion and for puns with a low level of aggressive content. In study 2, latent semantic incongruity was predictive of decreases in the emotional experience of "mirth" with repeated exposures in the context of a single experiment (but it did not predict changes in a more objective "cognitive appraisal" of humour). That is, puns with high (but not low) latent semantic incongruity significantly decreased in mirth with repeated exposures. In study 3, although semantic incongruity was positively associated with the number of elaborations provided by participants, semantic incongruity did not moderate the relationship between elaboration and humour. Humour produced by semantic incongruity appears to be independent from humour produced from elaboration. Semantic incongruity was associated with humour for puns with a low level of familiarity while elaboration was associated with humour for puns with a moderate level of familiarity.

The aforementioned novel findings regarding the function of semantic incongruity in humour from written puns were based on an operationalization of incongruity as the semantic dissimilarity between the two implied meanings of the pun. A similar approach was conducted by McHugh and Buchanan (2016); however, they did not explicitly examine whether semantic incongruity was associated with humour. Jared and Bainbridge (2017) found that the strongest predictor of humour was the semantic similarity between a key word from the context of the pun (e.g. "butcher") and the presented version of the homophone (e.g. "meat"; in the pun "the butcher was glad we could meat up"). The semantic *similarity* between a homophone and pun context was predictive of humour. Current results demonstrate that the semantic *dissimilarity* between the presented and alternate meanings of the pun is also predictive of humour (to follow

their example, this would be a comparison between "meat" and "meet"). They hypothesized that semantic similarity with a critical context word may have made the word play seem more unexpected or clever. Current results suggest that it is not only surprise in itself that is important to humour, but that humour is further based on the dissimilarity between the polysemous word with the alternative implied concept. That is, when puns with high semantic incongruity are more surprising, they should also be more humorous.

Current results replicate and expand on prior work by Hillson and Martin (1994), who demonstrated that the semantic dissimilarity between two concepts in artificial metaphor stimuli were predictive of humour ratings. Results from study 3 further suggest that semantic *similarity* may also be advantageous to elaboration on familiar content. Analogous results were obtained by Hillson and Martin were able to assess BOTH semantic dissimilarity and semantic similarities between concepts in artificial metaphor stimuli. They found that semantic dissimilarity was positively associated with humour regardless of the level of similarity in an item, but the most humorous metaphors were high in both semantic similarity and dissimilarity.

In study 3, latent semantic incongruity was not significantly associated with participant estimates of semantic relatedness of the two implied concepts in pun stimuli. This was assumed to be due to a ceiling effect for subjective estimates of semantic relatedness. However, further research comparing computational operationalizations of semantic relatedness to explore the validity of the current approach may be warranted.

5.1.2   Incongruity-Resolution

The incongruity-resolution model holds that humour should depend on the incongruity of an item, the complexity of problem solving necessary to explain incongruity, the time taken to resolve the incongruity, and the salience of the item's content (given emotional content such as aggression). A moderate level of comprehension challenge (i.e., stimuli should be neither too easy nor too difficult to understand) was hypothesized to be optimal for humour appreciation. Recall that critics of the incongruity-resolution model argue that incongruity-resolution confounds humour appreciation with humour comprehension.

According to this argument, a complete explanation of incongruities in humorous stimuli is not always possible, useful, or even desirable for humour appreciation (Forabosco, 1992, 2008; Ritchie, 2004). That is, you can completely understand a joke and yet not find it to be humorous, and you can find a joke to be humorous without completely understanding it.

Consistent with incongruity-resolution based predictions, latent semantic incongruity was positively associated with humour across all three studies. Although latent semantic incongruity was positively associated with humour, it was also associated with shorter comprehension durations (study 2). Semantic incongruity was found here to be associated with humour because it facilitates rapid and accurate (study 1) comprehension, rather than because it poses a comprehension challenge. In further support of incongruity-resolution, puns with high latent semantic incongruity were more humorous than puns with low latent semantic incongruity, but only for the first three exposures to an item (study 2) or for puns with a low level of familiarity from a prior occasion (study 1 and study 3). Although incongruity-resolution cannot clearly account for humour on repetition, shorter comprehension durations at first exposure were positively associated with humour at all subsequent repeated exposures to a pun. This finding indicates that incongruity-resolution at first exposure can continue to play a role in humour appreciation on repetition: initial comprehension difficulty appears to either create a good first impression of an item, or participants were simply able to recall the humour from comprehension difficulty at first exposure (as hypothesized by Suls, 1975).

Incongruity-resolution theory can not adequately account for the observed moderating effect of aggression on humour appreciation associated with latent semantic incongruity (study 1). Goldstein, Suls and Anthony's (1972; see also Suls, 1977) salience hypothesis argued that the purpose of aggressive content should be to facilitate resolution by drawing attention to important elements of a joke necessary to explain incongruous elements. Zillmann and Bryant (1980) made an analogous hypothesis (out of the assumption that all humour is inherently aggressive) that the purpose of incongruous content in jokes is to sanitize humour from aggressive content by allowing people to believe they are enjoying silly incongruities, rather than the aggressive content in itself.

Latent semantic incongruity and aggression did significantly interact but not as was predicted by either misattribution or the salience account. Semantic incongruity and aggression appeared to be mutually incompatible: there was an effect of aggression on humour only for puns with low semantic incongruity and there was an effect of semantic incongruity only for puns with low aggression (study 1). Taken together, these results also present the interesting possibility that aggressive humour may interfere with the extent to which people attend to semantic content. The participants may not actually have been attending to the semantic content (or at least the semantic humour) in puns with moderately aggressive content.

It is interesting to speculate whether aggression and semantic incongruity could interact productively for humour appreciation under different circumstances. The extent to which people enjoy and engage with moderately aggressive content should depend on individual preferences. Future research could therefore take participant preference for different kinds of humour styles into account (for example, see the humour styles questionnaire; Martin, Puhlik-Doris, Larsen, Gray, & Weir, 2003). Semantic incongruity may also be more important for enjoyment of humour from stimuli with higher levels of aggressive or offensive content (puns with highly aggressive or offensive content were not used in the currently reported studies), and aggressive content may be more important for stimuli with semantic incongruity for puns that are more difficult to comprehend (the puns used in studies 1-3 were relatively easy to understand).[12]

5.1.3   Comprehension-Elaboration

Building and expanding on predictions from incongruity-resolution, Wyer and Collins (1992) predicted that humour appreciation can be produced from two sources: from comprehension challenge (neither too easy nor too hard to understand the incongruities of humorous stimuli), and also from the linear quantity of elaboration on the implications of humorous stimuli *after* an item has been viewed and adequately understood. The first

---

[12] Study 1 mean clarity of understanding ratings for pun items (on a Likert-type scale from 0-6) = 4.59 SD = .76). Study 2 mean first exposure comprehension duration = 4310 ms, SD =1110, range = 2370 – 8280)

"comprehension" process is functionally identical to the incongruity-resolution process just described, and therefore this component of comprehension-elaboration makes analogous predictions. The second "elaboration" process was defined as the intentional generation of novel thoughts and features in association with the humorous aspects of stimuli beyond the time and effort necessary for comprehension. Comprehension-elaboration's greatest strength is arguably that it can account for humour appreciation from familiar stimuli. Familiar stimuli should remain humorous so long as people are able to generate novel humorous elaborations.

Wyer and Collins (1992) predicted that emotional content, such as aggression, should facilitate humour if it optimizes difficulty of comprehension (making it neither too easy nor too difficult to understand) or encourages elaboration. Aggression appears to serve as an independent source of humour; separate from, and incompatible with, humour from semantic incongruity. Studies 1 and 3 demonstrated that humour from aggression appears to be involved in both incongruity-comprehension and elaboration processes in humour appreciation. Moderately aggressive content in puns was associated with humour from both incongruity-resolution for novel stimuli, and with humour from elaboration for familiar stimuli. In study 1, moderate aggression was positively associated with humour only for puns with a low level of familiarity from a prior occasion. In study 3, the duration of time spent on the elaboration task was positively associated with humour from puns with a moderate level of aggressive content, but not for puns with a low level of aggressive content.

In study 2, the duration of time spent considering humour ratings for puns on the second and third exposures to an item was positively associated with humour. Puns with high latent semantic incongruity produced significantly higher mirth responses than puns with low latent semantic incongruity only for the same first three exposures. It was hypothesized that participants were either able to generate humour from re-interpreting the incongruities of the repeated pun or that they were using this time to generate novel elaborations on the content from these puns. Given that puns with high latent semantic incongruity produced significantly higher mirth responses than puns with low latent semantic incongruity only for the same first three exposures, I hypothesized that latent

semantic incongruity may be a meaningful indicator of the potential of an item for the generation of novel humorous insights.

Current findings were supportive of several comprehension-elaboration based predictions. Elaboration quantity and longer elaboration durations were positively associated with humour, but only for puns with a moderate level of familiarity from a prior occasion in study 3 (note that familiarity ratings had a comparatively smaller range in comparison to those from study 1).[13] Longer elaboration durations were also positively associated with humour for puns with aggressive content. There was also support for the hypothesis that elaboration *after* viewing and comprehending stimuli was important to humour appreciation. That is, participants produced a greater quantity of elaboration in a shorter duration of time and found the puns to be significantly more humorous when the elaboration task was completed after viewing and comprehending the puns (as opposed to elaboration prior to viewing the referent pun).

Latent semantic incongruity was significantly associated with elaboration quantity in study 3. However, there was no significant interaction between latent semantic incongruity and elaboration task variables in predicting humour ratings. That is, the relation between elaboration and humour was independent of latent semantic incongruity. This finding is inconsistent with the hypothesis that the positive association between rating duration and humour appreciation on second and third exposures was due to elaboration in study 2. Given the independence of latent semantic incongruity with humour from elaboration, it cannot be discounted that incongruity may have been associated with humour on second and third exposure because participants were able to reinterpret the pun's incongruity and generate new humour from resolution (as was proposed by Suls, 1972).

Wyer and Collins (1992) reflected that comprehension of incongruity in puns should be extremely easy and so it should be unlikely that humour from these items should be

---

[13] Study 1 overall mean familiarity for pun items = 2.57 *SD* = .62 (N = 300); study 3 mean familiarity = .83 *SD* = .40 (N = 100; both on a Likert-type scale from 0-6)

produce by differences in the challenge of comprehension. Their argument that puns should be easy to comprehend is further supported by results from Jared and Bainbridge (2017), who found longer fixation durations on the homophone in pun items (indicating that incongruity was recognized), but that there was no difference in total fixation times for pun sentences (indicating that the incongruities could be rapidly resolved). Current results indicate that fluency of comprehension was an important predictor of humour for puns with a low level of familiarity (study 1; and for first exposure to a pun in study 2). Given that latent semantic incongruity was positively associated with humour only for puns with a low level of familiarity from a prior occasion (study 1), it can therefore be hypothesized that semantic incongruity is important to humour produced by the "comprehension" process for novel or low familiarity items, whereas elaboration is important to humour appreciation produced from familiar puns. The study 2 finding that shorter durations of time necessary to understand each pun on first exposure was associated with humour ratings across all repeated exposures further indicates that humour from comprehension fluency and from elaboration are not necessary mutually exclusive: initial comprehension difficulty for novel stimuli may still impact humour appreciation on repetition. This is consistent with the comprehension-elaboration based hypothesis that the two processes can both actively play a role in humour appreciation.

In study 3, participant focus on semantic relatedness of the concepts in puns was experimentally manipulated. Participants were asked to rate all items for either the level of semantic dissimilarity or for the level of semantic similarity between the two implied concepts in puns. Participants who were assigned to rate semantic dissimilarities found the puns to be more humorous (as opposed to those focused on similarities), whereas in contrast participants who rated on semantic similarities produced a greater quantity of elaboration. Semantic dissimilarities were important to humour appreciation for unfamiliar stimuli, whereas this pattern of results poses the interesting possibility that semantic similarity (as opposed to dissimilarity) may play an important role in elaboration for humour appreciation for familiar stimuli. These results are consistent with findings from Hillson and Martin (1994) found that semantic dissimilarity (as they operationalized it) was associated with humour regardless of the level of semantic

similarity, but that the most humorous metaphors were high in both dissimilarity and similarity.

5.1.4 Fluency Account

Neither comprehension difficulty nor elaboration processes can clearly account for the overall importance of fluency in humour appreciation that was identified in the reported studies. By contrast, Topolinski's (2014) fluency account of humour appreciation proposed an accommodation to incongruity-resolution based models of humour appreciation that should be able to account for current findings. Topolinski hypothesized that humour appreciation should come from the positive feelings associated with promptness and surprising ease of insight (the "eureka!" experience) associated with understanding or explaining incongruities. Rather than the difficulty or problem solving challenge of explaining incongruities, the fluency of comprehension (easy, quick and productive comprehension) from incongruity-resolution should positively mediate the experience of humour.

On average, shorter durations of time spent processing written puns was positively associated with humour according to: pun identification accuracy and duration (study1), comprehension duration (study 2), and elaboration quantity and duration (study 3). Neither incongruity-resolution nor comprehension-elaboration can clearly account for these effects. There were several important exceptions to this pattern of results that should be considered. The biggest challenge to the fluency account is that the changes in humour appreciation with repeated exposures in study 2 were not associated with changes in comprehension duration. Fluency was also only important to humour from comprehension difficulty on first exposure (study 2) or from low familiarity stimuli (study 1). It is also problematic for the fluency account that, in study 2, longer durations of time spent providing humour ratings (rating duration) were predictive of humour at the second and third exposure to a pun. In study 3, longer elaboration durations were associated with humour for stimuli with a moderate level familiarity from a prior occasion and for puns with a moderate level of aggressive content. Here the only function of aggressive content that was relevant to the fluency account is that it may facilitate humour from elaboration duration. Taken together, these results indicate that the fluency

account only applies to humour from novel (or low familiarity) stimuli. Rapid processing is not *always* optimal for humour.

The fluency account may still yet be able to explain humour appreciation based on mere-exposure effects. The two-factor mere-exposure model (Berlyne, 1970) holds that when stimuli are repeated within a relatively short duration or a single context (as, for instance, was the case for the pun items presented in study 2), humour ratings should decrease with repetition due to a build-up of boredom or satiation with the topic. When stimuli are repeated over longer durations of time, humour appreciation should increase due to improved fluency. Consistent with this hypothesis, prior studies have found that humour ratings tend to decrease when repetition occurs within a single experiment (as in study 2; see also Goldstein, 1970b as cited in Suls, 1975; Pistole & Shor, 1979; Gavanski, 1986), whereas humour appreciation appears to persist (or even potentially increase as in studies 1-2) with familiarity from a non-specific prior occasion or when repetitions occur over longer durations (Schick, McGlynn & Woolam, 1972; Suls, 1975). In further support of the two factor model, Goldstein (1970b; as cited in Suls, 1975) found that the presence of sexual content (which may have reduced participant boredom) attenuated the statistically significant decreases in humour ratings over four repetitions of cartoons in the context of a single experiment. Zhang and Zinkhan (1991) found that humour ratings didn't significantly change over three repetitions of commercials embedded within a 30-minute TV program containing music videos. Repetitions of humorous stimuli embedded within entertaining content may have prevented decreases due to satiation and boredom. Further research investigating mere-exposure effects in humour appreciation across longer durations of time between repeated exposures is recommended.

5.1.5   Mirth and Cognitive Appraisal of Humour

The measurement by which humour is assessed proved to be important to proper interpretation of humour on repetition (consistent with predictions from Gavanski, 1986). Participant estimates of mirth (the current emotional experience to humour in response to stimuli) but not participant "cognitive appraisal" of humour from latent semantic incongruity (the more objective overall assessment of the quality of the items themselves) significantly decreased over repeated exposures to puns in study 2. In study 2, cognitive

appraisal ratings were found to be significantly higher than mirth ratings. The participants likely recognized that their current emotional experience of humour may be lower than the more objective overall quality of humorous stimuli. In study 3, mirth and cognitive appraisal ratings of humour were highly correlated and there were no meaningful distinctions between the two scales. The null difference between the two humour scales in study 3 may be due to the relatively low familiarity ratings (compared with study 1), or due to mere-exposure (over repetitions, satiation or boredom might reduce the sense of mirth reaction but not an appreciation of the overall funniness of the joke). The distinct pattern of results for mirth and cognitive appraisal in study 2 (in which repetitions occur within the context of a single one hour experiment) nevertheless suggest that prior studies which did not distinguish between assessments of item quality versus participant emotional reaction to humour may have underestimated decreases in humour with repeated exposures (for example, this may have been the case for Zhang & Zinkhan, 1991; Suls, 1975; or due to familiar content as in Schick, McGlynn & Woolam, 1972). Future studies should further explore the conditions under which mirth and cognitive appraisal of humour diverge.

5.1.6   Further Considerations

The variables employed here were often abstract operationalizations of constructs discussed in theories of humour appreciation. Nonetheless, the current results had considerable face validity: they were associated with humour ratings as would be predicted by explanatory models of humour appreciation. Future research should further investigate these variables through additional converging methods to strengthen confidence of current findings. However, the current results should generalize beyond puns to other forms of humorous stimuli such as metaphors, jokes or cartoons. Further research extending current findings with a broader and more representative range of possible types of pun items is also warranted (e.g. with a greater range of aggressive or offensive content).

There may have been insufficient variability in comprehension challenge of written pun stimuli given the undergraduate participants that were sampled in the current research. There was a high degree of reliability between participant estimates of humour ratings for

the pun stimuli, but it is unclear if this was due to the objective quality of the pun items or due to the relatively homogenous sample of participants and the easily understood pun stimuli. One cannot discount that the hypothesized non-linear effect of comprehension challenge on humour ratings may yet be identified with a broader range of participants, such as with children (as in Zigler, Levine & Gould, 1967; McGhee, 1976) or with adults with less proficiency in the language (for example, as with novice learners of English as a second language).

Across studies, several assumptions and approximations where made for the time duration variables and therefore further replication with greater experimental control may be warranted. For example, in study 2 it was assumed that comprehension duration, the time it takes participants to "get" the humour in each pun, was representative of the humour produced from comprehension of semantic incongruities. It is unclear whether comprehension duration on repeated exposures was unassociated with humour on repeated trials because they could be understood more rapidly with repetition, or if participants simply immediately pressed the "got it" button upon recognizing an item as familiar. It was hypothesized that participants may have been re-considering humour ratings on second and third exposure to a pun and that they may have been generating novel elaborations during this "rating duration" period. This assumption was not directly tested, but study 3 results regarding elaboration in humour were consistent with this interpretation. Future studies should use more precise approaches to examine the online comprehension and appreciation of humour from semantic incongruities (for example using eye tracking or event-related potential methodology).

In study 3, the function of elaboration in humour was operationalized and tested using a task which asked participants to list the concepts which come to mind when considering the two implied meanings in pun stimuli. In natural settings, it is unlikely that participants elaborate on humour from puns in this fashion. Results from the elaboration task supported the core assumption from the comprehension-elaboration model (Wyer & Collins, 1992) that puns should be more humorous for stimuli that can more easily and productively generate associated concepts. However, it is unclear if rapid elaboration task performance for novel stimuli was associated with humour for unfamiliar puns due to

fluency effects, of if spending a longer duration of time on this artificial task made the puns less fun. The findings from study 3 should replicate with a more naturalistic elaboration on familiar humorous stimuli. Wyer and Collins (1992) noted that elaboration does not have to be personally generated. Therefore, guiding participants to new ideas, concepts, or implications of familiar stimuli should generate analogous effects. As an anecdotal example, comedians often generate humour not only with the build-up and punchline of a joke, but also from an exploration of consequences or with a re-interpretation and elaboration on familiar stimuli. Providing participants with more naturalistic elaborations (rather than requiring them to generate them on their own) should prove to be an effective method of addressing this concern.

5.1.7   Conclusion: Evaluation of Humour in Puns

Studies 1-3 operationalized key concepts from the humor appreciation literature for experimental study (such as semantic incongruity or resolution) and tested key predictions from major explanatory models of humour appreciation. The use of written puns as stimuli in the present research proved to be ideal to investigate predictions from these models. On review, the comprehension-elaboration model (Wyer & Collins, 1992) appears to do the best job at accounting for current findings (if accommodations are made for the moderating influences of familiarity, aggression, and fluency). Humour appreciation appears to be produced from semantic incongruity (*or* aggression) and comprehension fluency for puns with a low level of familiarity, and from longer elaboration for moderately familiar puns. Across the studies conducted for this dissertation, far from being the lowest form of humour, participants indicated that puns were, on average, moderately humorous.

The focus here was on evaluating cognitive theories of humour, however, as a parting point, the data presented here indicates that if you want to be humorous you should attempt to be incongruous and, on some level, include familiar content. It would also be beneficial to include moderately aggressive content given that it improved to be important to humour regardless of the level of familiarity in a pun. The best novel puns are semantically incongruous, can be rapidly understood on first exposure, and can produce a greater quantity of elaboration over repeated exposures.

References

Amir, O., Biederman, I., Wang, Z., & Xu, X. (2015). Ha Ha! Versus Aha! A direct comparison of humor to nonhumorous insight for determining the neural correlates of mirth. *Cerebral Cortex, 25*(5), 1405-1413.

Attardo, S. (1994). *Linguistic theories of humor*. Hawthorne, NY: Mouton de Gruyter.

Attardo, S. (1997). The semantic foundations of cognitive theories of humor. *Humor: International Journal of Humor Research*, *10*(4), 395–420.

Attardo, S., & Raskin, V. (1991). Script theory revis(it)ed: Joke similarity and joke representation model. *Humor: International Journal of Humor Research*, *4*(3–4), 293–347.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York, NY: McGraw-Hill.

Berlyne, D. E. (1970). Novelty, complexity and hedonic value. *Perception & Psychophysics*, *8*, 279–286.

Bornstein, R. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, *106*, 265–289.

Bornstein, R. F., & D'Agostino, P. R. (1992). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology*, *63*, 545-552.

Bornstein, R. F., & D'Agostino, P. R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/attributional model of the mere exposure effect. *Social Cognition*, *12*(2), 103-128.

Bornstein, R. F., Kale, A. R., & Cornell, K. R. (1990). Boredom as a limiting condition on the mere exposure effect. *Journal of Personality and Social Psychology*, *58*(5), 791-800.

Dawson, J. F., & Richter, A. W. (2006). Probing three-way interactions in moderated multiple regression: Development and application of a slope difference test. *Journal of Applied Psychology*, *91*, 917-926.

Dictionary.com. Definitions Retrieved August, 2014 from https://www.dictionary.com/.

Dryden, J. (1672). Defence of the epilogue; or, An essay on the dramatic poetry of the last age. In Appended to *The conquest of Granada by the Spaniards in two parts* (pp. 161–175). London, England: TN for Henry Herringman. Retrieved from http://name.umdl.umich.edu/a36610.0001.001

Duchacek, O. (1970) Les jeux de mots du point de vue linguistique. *Beitriige zur romanischen Philologie*. *9*: 1. 107-11 7.

Durda, K., & Buchanan, L. (2006). WordMine2 [Online] Available: http://web2.uwindsor.ca/wordmine.

Durda, K., & Buchanan, L. (2008). WINDSORS: Windsor improved norms of distance and similarity of representations of semantics. *Behavior Research Methods*, *40*(3), 705-712.

Forabosco, G. (1992). Cognitive aspects of the humor process: The concept of incongruity. *Humor: International Journal of Humor Research*, *5*(1-2), 45-68.

Forabosco, G. (2008). Is the concept of incongruity still a useful construct for the advancement of humor research? *Lodz Papers in Pragmatics*. 45-62.

Forabosco, O. (1992). Cognitive aspects of the humor process: The concept of incongruity. *Humor: International Journal of Humor Research*, *5*, 45-68.

Gavanski, I. (1986). Differential sensitivity of humor ratings and mirth responses to cognitive and affective components of the humor response. *Journal of Personality & Social Psychology*, *51*(1), 209–214.

Gelb, B. D. & Zinkhan, G. M. (1985). The effect of repetition on humor in a radio advertising study. *Journal of Advertising*. *14*, 13-20.

Goldstein, J. H. (1970). Humor appreciation and time to respond. *Psychological Reports*, *27*(2), 445-446.

Goldstein, J. H. (1970). Repetition, motive arousal, and humor appreciation. *Journal of Experimental Research in Personality, 4*(2), 90-94.

Goldstein, J. H., Suls, J. M., & Anthony, S. (1972). Enjoyment of specific types of humor content: Motivation or salience? In J. H. Goldstein & P. E. McGhee (Eds.), *The psychology of humor: Theoretical perspectives and empirical issues* (pp. 159–171). New York: Academic Press.

Gruner, C. R. (1997). *The game of humor: A comprehensive theory of why we laugh*. New Brunswick, NJ: Transaction Publishers.

Günther, F., Dudschig C., & Kaup, B. (2016). Latent semantic analysis cosines as a cognitive similarity measure: Evidence from priming studies, *The Quarterly Journal of Experimental Psychology*, *69*:4, 626-653.

Herzog, T. R., & Bush, B. A. (1994). The prediction of preference for sick humor. *Humor: International Journal of Humor Research*, *7*(4), 323–340.

Herzog, T. R., & Karafa, J. A. (1998). Preferences for sick versus nonsick humor. *Humor: International Journal of Humor Research*, *11*(3), 291–312.

Hillson, T. R., & Martin, R. A. (1994). What's so funny about that?: The domains-interaction approach as a model of incongruity and resolution in humor. *Motivation & Emotion*, *18*(1), 1–29.

Holland, R. W., Hendriks, M., & Aarts, H. (2005). Smells like clean spirit: nonconscious effects of scent on cognition and behavior. *Psychological Science*, *16*(9), 689-693.

Jacoby, L. L., & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin*, *13*, 314– 336.

Jacoby, L. L., Toth, J. P., Lindsay, D. S., & Debner, J. A. (1992). Lectures for a layperson: Methods for revealing unconscious processes. In R. F. Bornstein & T. S. Pittman (Eds.), *Perception without awareness* (pp. 81–120). New York, NY: Guilford Press.

Jared, D., & Bainbridge, S. (2017). Reading homophone puns: Evidence from eye tracking. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 71*(1), 2-13.

Kao, J. T., Levy, R., & Goodman, N. D. (2016). A computational model of linguistic humor in puns. *Cognitive Science, 40*(5), 1270-1285.

Kenny, D. T. (1955). The contingency of humor appreciation on the stimulus-confirmation of joke-ending expectations. *Journal of Abnormal & Social Psychology*, *51*, 644–648.

Koestler, A. (1964). *The act of creation*. London: Hutchinson.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211-240.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, *25*, 259-284.

Martin, R. A. (2007). *The psychology of humor: An integrative approach*. Amsterdam, Netherlands: Elsevier.

Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. Journal of Research in Personality, *37*(1), 48–75.

McGhee, P. E. (1976). Children's appreciation of humor: A test of the cognitive congruency principle. *Child Development*, *47*(2), 420–426.

McHugh, T., & Buchanan, L. (2016) Pun processing from a psycholinguistic perspective: Introducing the Model of Psycholinguistic Hemispheric Incongruity Laughter (M.PHIL), *Laterality: Asymmetries of Body, Brain and Cognition*, *21*:4-6, 455-483.

Merriam-Webster.com. Definitions Retrieved August, 2014, from https://www.merriam-webster.com/dictionary/

Montoya, R. M., Horton, R. S., Vevea, J. L., Citkowicz, M., & Lauber, E. A. (2017). A re-examination of the mere exposure effect: The influence of repeated exposure on recognition, familiarity, and liking. *Psychological Bulletin*, *143*(5), 459-498.

O'Shannon, D. (2012). *What Are You Laughing at? A Comprehensive Guide to the Comedic Event*. New York, NY: Bloomsbury Academic.

Pien, D., & Rothbart, M. K. (1977). Measuring effects of incongruity and resolution in children's humor. In A. J. Chapman & H. C. Foot (Eds.), *It's a funny thing, humour* (pp. 211-213). Oxford: Pergamon Press.

Pistole, D. D., & Shor, R. E. (1979). A multivariate study of the effect of repetition on humor appreciation as qualified by two social influence factors. *Journal of General Psychology*, *100*(1), 43-51.

Pollio, H. R., & Mers, R. W. (1974). Predictability and the appreciation of comedy. *Bulletin of the Psychonomic Society*, *4*(4-A), 229–232.

Punoftheday.com. Public domain puns retrieved August, 2014 from https://www.punoftheday.com/

Punsandjokes.com. Public domain puns retrieved August, 2014 from http://www.punsandjokes.com/.

Ritchie, G. (1999). Developing the incongruity-resolution theory. In *Proceedings of the AISB Symposium on Creative Language: Stories and Humour* (pp. 78–85). East Sussex, UK: Society for AISB.

Ritchie, G. (2004). *The linguistic analysis of jokes*. London: Routledge.

Ritchie, G. (2009). Variants of incongruity resolution. *Journal of Literary Theory*, *3*(2), 313–332.

Ruch, W. (1981). Humor and personality: A three-modal analysis. *Zeitschrift fur Differentielle und Diagnostische Psychologie*, *2*, 253–273.

Ruch, W. (1984). Conservatism and the appreciation of humor. *Zeitschrift fur Differentielle und Diagnostische Psychologie*, *5*, 221–245.

Ruch, W. (1992). Assessment of appreciation of humor: Studies with the 3 WD Humor Test. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in personality assessment* (Vol. *9*, pp. 27–75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ruch, W., & Hehl, F. J. (1998). A two-mode model of humor appreciation: Its relation to aesthetic appreciation and simplicity-complexity of personality. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic* (pp. 109–142). Berlin, Germany: Walter de Gruyter.

Schick, C., McGlynn, R. P., & Woolam, D. (1972). Perception of cartoon humor as a function of familiarity and anxiety level. *Journal of Personality and Social Psychology*, *24*(1), 22-25.

Schultz, T. R. (1972). The role of incongruity and resolution in children's appreciation of cartoon humor. *Journal of Experimental Child Psychology*, *13*, 456-477.

Shakespeare, W. (2015). The Complete Works of William Shakespeare. Accessed August, 2018. http://shakespeare.mit.edu.

Sheridan, H., Reingold, E. M., & Daneman, M. (2009). Using puns to study contextual influences on lexical ambiguity resolution: Evidence from eye movements. *Psychonomic Bulletin & Review, 16*(5), 875-881.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059-1074.

Stang, D. J. (1973). Six theories of exposure and affect. *Catalog of Selected Documents in Psychology*, *3*, 126.

Suls, J. M. (1972). A two-stage model for the appreciation of jokes and cartoons: An information-processing analysis. In J. H. Goldstein & P. E. McGhee (Eds.) *The psychology of humor: Theoretical perspectives and empirical issues* (pp. 81-100). New York: Academic Press.

Suls, J. M. (1975). The role of familiarity in the appreciation of humor. *Journal of Personality*, *43*(2), 335-345.

Suls, J. M. (1977). Cognitive and disparagement theories of humour: A theoretical and empirical synthesis. In A. J. Chapman & H. C. Foot (Eds.), *It's a funny thing, humour* (pp. 41–45) Oxford: Pergamon Press.

Suls, J. M. (1983). Cognitive processes in humor appreciation. In P. E. McGhee & 3. H. Goldstein (Eds.) *Handbook of humor research*, Vol. 1. Basic Issues (pp. 39-57). New York: Springer-Verlag.

Suls, J.M. (1983). Cognitive processes in humor appreciation. In P.E. McGee & J.H. Goldstein (Eds.), *Handbook of humor research, Vol. 1: Basic issues*, 39-57. New York: Springer-Verlag.

Topolinski, S. (2014). A processing fluency-account of funniness: Running gags and spoiling punchlines. *Cognition & Emotion*. *28*, 811–820.

Tourangeau, R., & Sternberg, R. J. (1982). Understanding and appreciating metaphors. *Cognition*, *11*, 203-244.

Trick, L., & Katz, A. N. (1986). The domain interaction approach to metaphor processing: Relating individual differences and metaphor characteristics. *Metaphor and Symbolic Activity*, *1*, 203-244.

Utsumi, A. (2005). The role of feature emergence in metaphor appreciation, *Metaphor and Symbol*, *20:3*, 151-172.

Wyer, R. S., & Collins, J. E. (1992). A theory of humor elicitation. *Psychological Review*, *99*(4), 663–688.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*, 1-27.

Zhang, Y., & Zinkhan G.M. (1991). Humor in Television Advertising: The Effects of Repetition and Social Setting. In *NA - Advances in Consumer Research Volume*

18, eds. Rebecca H. Holman and Michael R. Solomon, Provo, UT: Association for Consumer Research, Pages: 813-818.

Zigler, E., Levine, J., & Gould, L. (1966). Cognitive processes in the development of children's appreciation of humor. *Child Development*, *37*(3), 507-518.

Zigler, E., Levine, J., & Gould, L. (1967). Cognitive challenge as a factor in children's humor appreciation. *Journal of Personality and Social Psychology*, *6*(3), 332-336.

Zillmann, D., & Bryant, J. (1974). Retaliatory equity as a factor in humor appreciation. *Journal of Experimental Social Psychology*, *10*(5), 480–488.

Zillmann, D., & Bryant, J. (1980). Misattribution theory of tendentious humor. *Journal of Experimental Social Psychology*, *16*(2), 146–160.

Zillmann, D., & Cantor, J. R. (1976). A disposition theory of humour and mirth. In A. J. Chapman & H. C. Foot (Eds.), *Humor and laughter: Theory, research, and applications* (pp. 93–115). London: John Wiley & Sons.

Zillmann, D., Bryant, J., & Cantor, J. R. (1974). Brutality of assault in political cartoons affecting humor appreciation. *Journal of Research in Personality*, *7*(4), 334-345.

Appendix A: Complete list of puns and control items and frequency estimates.

Note: Study 2a refers to items used as non-repeated filler puns, 2b refers to repetition items, 2c refers to study 2 lure items. Item types include homophones (HP), homographs (HG), rhymes (RH), control items (CI) and lures (LU). Control item and pun versions would not both appear within the same experimental list. Orthographic frequency (OFREQ; N=281 of 300 puns) and phonological frequency (PFREQ; N=271) estimates from wordmine2 are reported for study 1 pun items based on wordmine2 estimates for the ambiguous word where relevant entries were available for the ambiguous word as spelled or according to the dominant implied meaning.

	Pun	Item type	study #	Ambiguous word	Ofreq	Pfreq
1	A backward poet writes inverse.	HP	1	inverse	2.82	2.02
2	A baker stopped making donuts after he got tired of the hole thing.	HP	1, 2a, 3	hole	56.68	0.38
3	I finally got rid of that nasty electrical charge I've been carrying. I'm ex-static!	HP	1	ecstatic	3.14	1.58
4	A botanist-turned-prize fighter was penalized for aloe blow.	HP	1, 2b	aloe	0.43	1.11
5	A cardboard belt would be a waist of paper.	HP	1, 2b, 3	waist	23.62	1.40
6	A carpenter must have been here. I sawdust.	HP	1	sawdust	2.07	0.17
7	A chicken was murdered yesterday. The investigator thinks there was some fowl play involved.	HP	1, 2b, 3	fowl	5.63	2.40
8	A fisherman hated fish and chips but he didn't tell a sole.	HP	1	sole	25.66	2.52
9	A good insurance company knows how to handle acclaim.	HP	1	acclaim	1.67	0.00
10	A hawk sat atop a church because it was a bird of pray.	HP	1, 2a, 3	pray	42.73	2.49
11	A lawyer-turned-cook is a sue chef.	HP	1	sue	24.98	0.94
12	A man who wanted to sing in church was wondering if he should inquire.	HP	1	inquire	10.68	2.52
13	A man with no pennies got into senseless trouble.	HP	1	senseless	5.75	2.76
14	A meteor just crashed into Russia ... no comet.	HP	1	comet	4.15	1.22
15	A young girl in charge of her tribe would be called little miss-chief.	HP	1	mischief	12.84	3.97
16	After hearing the case of the woman who folded her clothes wrong, the jury had no choice but to hanger.	HP	1	hanger	1.21	0.03
17	After taking the elevator to the top floor I felt very up-lifted.	HP	1	uplifted	4.93	3.58

18	All the waterfowl kept their eyes closed except for one. He was a Peking Duck.	HP	1	peking	2.86	1.06
19	An electrician knows watt is important.	HP	1, 2a, 3	watt	3.87	3.89
20	An English teacher, who was dreadfully afraid of insects, while on a picnic screamed like a little girl when he saw there was an antonym.	HP	1	antonym	0.09	
21	Are dog biscuits made of collie flour?	HP	1	cauliflower	0.93	2.23
22	Are Philosophy papers graded with Marx out of ten?	HP	1	marx	7.00	1.15
23	Atheism is a non-prophet organization.	HP	1, 2a, 3	prophet	16.35	0.40
24	Bakers trade bread recipes on a knead to know basis.	HP	1, 2a, 3	knead	0.68	1.02
25	Being a baker is hard, you've probably got to take on many rolls.	HP	1, 2a, 3	rolls	13.35	1.59
26	Being a poet in prison surely has its prose and cons.	HP	1	prose	10.22	0.97
27	Bought an apple, took a byte out of it.	HP	1	byte	14.99	2.04
28	Even during a zombie apocalypse, I'd still chews you.	HP	1, 2b, 3	chews	0.34	0.34
29	I agreed to watch my neighbor's dog, but only if it didn't scratch me. It's in the clause.	HP	1, 2a, 3	clause	17.84	1.17
30	I bet the butcher $50 that he couldn't reach the meat off the top shelf. He said, 'no, the steaks are too high.'	HP	1, 2b, 3	steaks	1.20	1.39
31	I like to hang my rugs on the clothesline. To air is human.	HP	1	air	282.57	2.90
32	I need to do my philosophy homework but I just Kant.	HP	1			
33	My supply of pants is being depleated.	HP	1	depleted	2.01	1.38
34	A bacteria walked into a bar and the bartender said, 'We don't serve bacteria in this place.' The bacteria said, 'But I work here, I'm staph.'	HP	1			
35	At breakfast, the hacker downloaded cornflakes via his cereal port.	HP	1, 2b, 3	cereal	2.35	0.97
36	Bugs have very diverse religious views, because they are all in sects.	HP	1	insects	16.44	0.41
37	Cannibals like to meat people.	HP	1, 2a, 3	meat	43.02	0.64
38	Chemistry jokes may be old and dead, but I just can't seem to Barium.	HP	1	barium	0.67	1.00
39	Coming up with cheese puns should be a bries.	HP	1	brie	0.38	1.18

40	Correspondence citing farm machinery defects would be a Deere John letter.	HP	1	dear	257.95	2.95
41	Could we really blame the hurricane for all of the broken glass? Realistically, how much can a wind owe?	HP	1	window	168.64	99.65
42	Did you hear about the fire at the shoe factory? A thousand soles were burned.	HP	1, 2a, 3	soles	3.07	0.00
43	Did you hear about the injured vegetable? Some say he got beet.	HP	1, 2a, 3	beet	1.26	0.00
44	Did you hear about the Italian Chef that died? He pasta way.	HP	1	pasta	2.32	1.58
45	Do hotel managers get board with their jobs?	HP	1, 2a, 3	board	118.86	1.18
46	Do the people who climb the world's highest mountain ever rest?	HP	1			
47	Do you know how winter coats are insulated? They are down loaded.	HP	1	down	1234.04	2.30
48	Do you know why Frequency cannot love any more? Cause it still hertz.	HP	1	hertz	0.84	0.55
49	Does it burn when you pee? It sure sounds like urine trouble to me.	HP	1	urine	4.50	2.50
50	Drove my car into a tree once and finally figured out how a Mercedes bends.	HP	1	bends	3.70	0.90
51	Eating should never make you sad, unless it is a mourning meal.	HP	1	mourning	9.77	0.00
52	Even crazy people know that you should wear hearing protection in high noise areas. That's ear rational.	HP	1	irrational	4.05	1.07
53	Have you heard the story about my arm? It's pretty humerus.	HP	1, 2a, 3	humerus	0.40	1.57
54	He avoided funerals because he was not a mourning person.	HP	1, 2a, 3	mourning	9.77	0.00
55	He rose through the ranks of the International Corn Growers association, eventually becoming a kernel.	HP	1, 2a, 3	kernel	2.76	1.58
56	Hey are you a bank? Because you need to leave me a loan.	HP	1	loan	20.39	1.47
57	Hotel owners usually have suite dreams.	HP	1, 2a, 3	suite	11.71	0.48
58	How did I get out of Iraq? Iran.	HP	1	iran	8.04	2.11
59	How do spacemen add more protein to their diet? They make it meteor.	HP	1	meteor	1.51	1.19
60	How do you make antifreeze? Steal her blanket.	HP	1	antifreeze	0.05	
61	How do you organize a space party? You planet.	HP	1	planet	16.24	2.49
62	How do you wake up Lady Gaga? Poke her face.	HP	1			
63	How does a man on the moon cut his hair? Eclipse it.	HP	1	eclipse	3.66	0.01

64	How many female priests does the Catholic church have? Nun.	HP	1, 2a, 3	nun	5.60	1.09
65	How many tickles does it take to make an octopus laugh? Ten-tickles.	HP	1	tentacles	1.85	1.30
66	I always prayed before my trigonometry tests. I was hoping for a sine from above.	HP	1, 2a, 3	sine	2.06	0.21
67	I bought me some of those new paper shirts. I don't like them -- they're tearable.	HP	1	terrible	84.33	3.26
68	I can never wear glasses. They make me see-sick.	HP	1	see	1404.96	2.80
69	I changed my iPod's name to Titanic and now its syncing.	HP	1	sync	0.32	0.78
70	I considered going into the ministry but I didn't have an altar ego.	HP	1, 2a, 3	altar	19.63	1.85
71	I didn't want to buy leather shoes, but eventually I was suede.	HP	1, 2b, 3	suede	1.28	0.28
72	I donated my blood and they told me I was Type-A. Apparently it was a Type-O.	HP	1			
73	I found a dried grape with my dried cranberries. Going to start raisin hell.	HP	1, 2b, 3	raisin	0.86	1.59
74	I had a boyfriend with a glass eye, he was shattered when eye dumped him.	HP	1, 2a	eye	161.17	3.46
75	I had a little bird, her name was enza, I opened up the window and influenza.	HP	1	influenza	1.30	1.16
76	I hate people who make bad puns, they should be pun-ished.	HP	1	punished	12.70	0.28
77	I hate the price of candy at the movie theater. They're always raisinette.	HP	1			
78	I hear that strangers are living in your basement. Of course, these are only roomers.	HP	1	roomers	0.09	
79	I just can't get enough of stories with female protagonists. I guess you could say I'm a heroine addict.	HP	1, 2a, 3	heroine	6.37	1.91
80	I just read about a fellow who rode on the ferris wheel longer than anyone else. It is in the Guinness Book of Whirled Records.	HP	1, 2a, 3	whirled	7.36	0.00
81	I love the internet. It's like the wifi never had.	HP	1			
82	I made a batch of fish eye soup, it should sea me through the week.	HP	1, 2b, 3	sea	224.71	1.46
83	I met a man who loves eating couches. I think he has a suite tooth.	HP	1	suite	11.71	0.48
84	I only think I am once in a while. Cogito ergo some.	HP	1	some	1787.50	2.00

85	I ordered whole wheat toast but it tasted funny. I think something was awry.	HP	1	awry	1.66	0.02
86	I saw a beaver movie last night, it was the best dam movie I've ever seen.	HP	1, 2a, 3	dam	7.81	0.16
87	I shot a man in paintball just to watch him dye.	HP	1, 2a, 3	dye	4.36	2.40
88	I studied a long time to become a doctor, but I didn't have any patience.	HP	1	patience	28.31	4.56
89	I tried working in a bakery, but was told I wasn't 'bread' for it.	HP	1, 2b, 3	bread	59.08	1.91
90	I used to do balance and rotations at an auto shop. I felt like I was just spinning my wheels and decided to retire.	HP	1	retire	11.99	1.59
91	I used to do rock climbing as a youth, but I was much boulder back then.	HP	1	boulder	3.91	0.06
92	I was going to study the work of Sigmund Freud, but I was too Jung to understand it.	HP	1			
93	I went to a seafood disco last week and pulled a mussel.	HP	1, 2a, 3	mussel	0.55	2.07
94	Some daze I just can't seem to focus.	HP	1, 2a	daze	1.51	1.58
95	If you lose your hearing, is it ear replaceable?	HP	1	ear	62.37	1.97
96	Lettuce continue with the bad jokes.	HP	1	lettuce	3.37	2.75
97	My friend talked me out of eating an unhealthy meal. I saw the error of my weighs.	HP	1, 2a, 3	weighs	2.88	1.79
98	Old actors never die, they just drop a part.	HP	1	apart	75.46	2.32
99	Seven days without pizza makes one weak.	HP	1, 2b, 3	weak	57.53	2.27
100	A ship's captain is a sails manager.	HP	2a			
101	I'm trying to find a rope tying class, should I look for a knot for profit organization?	HP	2a			
102	If children were allowed to dig for coal, would they still be miners?	HP	2a			
103	If everybody loves Christmas so much…why don't they all just Merry Christmas?	HP	2a			
104	In the morning a lawyer walked on his lawn and experienced the dew process.	HP	2a			
105	It was an emotional wedding, even the cake was in tiers.	HP	2a			
106	Sailing is like oil drilling because they're both crewed businesses.	HP	2a			

107	Today is a kick back and do nothing day. I'm auditioning for American Idle.	HP	2a			
108	What do you call an illegally parked frog? Toad.	HP	2a			
109	When Bambi's mom died, a lot of people lost a deer friend.	HP	2a			
110	When the gunman walked in, he turned the store into a flee market.	HP	2a			
111	If you ride a bus you have to pay your fare share.	HP	2a			
112	The indecisive rower couldn't choose either oar.	HP	2a			
113	The magician got so mad he pulled his hare out.	HP	2a			
114	Don't expect to eat something fancy when you're flying because it's plane food.	HP	2a			
115	Funniness and cleverness have always been two notable factors for rating puns, but the third has groan in significance.	HP	2a			
116	Herb gardeners who work extra get thyme and a half.	HP	2a			
117	A fight over love and money would be duel purpose.	HP	2a			
118	The mime wanted to say something, but he wasn't aloud.	HP	2a			
119	A baby chicken has a hard time coming out of its shell.	HG	1	shell	28.52	0.87
120	A bad shoemaker's assistant was given the boot.	HG	1, 2a, 3	boot	14.83	0.76
121	A baseball player can sell himself to a new team if he has a good pitch.	HG	1, 2a, 3	pitch	22.26	1.16
122	A boiled egg in the morning is hard to beat.	HG	1, 2a	beat	65.05	2.31
123	A book called 'Current Trends in Wiring your House' turned out to be a shocking failure.	HG	1, 2b, 3	shock	45.00	2.55
124	A carpenter sat on his drill and was bored to tears.	HG	1, 2a, 3	bore	42.14	3.16
125	A doctor who became a bartender was always giving out shots!	HG	1, 2a, 3	shot	100.68	3.57
126	A dog gave birth to puppies near the road and was ticketed for littering.	HG	1	litter	7.94	0.01
127	I visit the dentist frequently. So I know the drill.	HG	1	drill	8.73	0.06
128	A fisherman tried boxing, but he only threw hooks.	HG	1, 2b, 3	hook	16.42	1.60
129	A flat rate is the monthly rent for an apartment.	HG	1, 2a, 3	flat	72.39	2.47
130	A gardener who moved back to his home town rediscovered his roots.	HG	1, 2a	root	32.28	0.14

131	Daylight savings is really brightening my mood.	HG	1, 2a, 3	bright	100.75	2.53
132	A guy walks into a bar with jumper cables. The bartender says, 'You can come in, but don't start anything!'	HG	1, 2a, 3	start	176.71	1.78
133	A horse walks into a bar. The bartender says, 'So, why the long face?'	HG	1	long	896.36	2.54
134	I tried looking for gold, but it didn't pan out.	HG	1	pan	21.36	1.33
135	A king measures his line with a ruler.	HG	1, 2b, 3	ruler	11.18	2.24
136	A man was hit by a liquor truck. Looks like this time the drinks were on him	HG	1	on	6750.97	2.81
137	I tried wearing tight jeans, but I can never pull them off.	HG	1, 2b, 3	pull	45.73	3.58
138	A new batter joined the baseball team, and he was a real hit.	HG	1	hit	71.13	2.67
139	I tried to get a job at a casino but they didn't have a slot for me.	HG	1	slot	8.15	1.96
140	An egg was late for work, he scrambled to get ready.	HG	1, 2a, 3	scramble	3.96	0.00
141	Although I like people who are very open, I will never become a surgeon.	HG	1	open	338.04	2.34
142	Aliens can easily understand each other because their language is universal.	HG	1, 2a	universal	38.10	3.67
143	After the test drive, the car salesman drove home his point.	HG	1	drove	55.44	2.25
144	After 5 years with the same chiropractor, I moved and had to change doctors. It was quite an adjustment.	HG	1, 2a, 3	adjustment	8.64	0.05
145	After he ate the duck, the alligator got a little down in the mouth.	HG	1	down	1234.04	2.30
146	After he invented the light bulb, people saw Edison in a new light.	HG	1	light	377.06	2.50
147	After Junior swallowed the watch he had to wait to pass the time.	HG	1, 2a, 3	pass	131.84	2.71
148	After my ear operation I feel sound.	HG	1	sound	182.31	2.26
149	After periodic doubts about his vocational calling, the young chemistry teacher concluded he was out of his element.	HG	1	element	45.90	2.30
150	Artists are colorful people who draw on their emotions.	HG	1, 2a	draw	72.11	1.58
151	AT&T and T-mobile got married, I heard the reception was terrible.	HG	1, 2b, 3	reception	21.86	3.77
152	Did you hear about that new drug that makes people angry? It's all the rage now.	HG	1, 2a, 3	rage	31.60	3.06
153	A dog not only has a fur coat but also pants.	HG	1	pant	1.31	1.01

154	A group called the Balloons sings pop music.	HG	1, 2a, 3	pop	19.42	0.10
155	A hot dog vendor didn't relish his job.	HG	1, 2a	relish	5.77	1.98
156	A murderer started a business and made a killing.	HG	1, 2b, 3	kill	71.36	2.54
157	A prisoner's favorite punctuation mark is the period. It marks the end of his sentence.	HG	1, 2a, 3	sentence	46.95	1.61
158	At a hearing aid center: 'Let us give you some sound advice.'	HG	1, 2b, 3	sound	182.31	2.26
159	Be true to your teeth, or they will be false to you.	HG	1	false	53.67	3.13
160	Cartoonist found dead in home. Details are sketchy.	HG	1, 2a, 3	sketch	14.04	2.78
161	Coaches usually have a goal in mind.	HG	1	goal	49.40	1.51
162	Contemplating my imminent root canal procedure was deeply unnerving.	HG	1	nerve	19.08	0.27
163	Continually discovering new viruses can become a strain.	HG	1, 2a	strain	30.08	2.24
164	Dead batteries are typically sold free of charge.	HG	1, 2a, 3	charge	100.71	3.23
165	Did you hear about the crab in financial difficulty? It was starting to feel the pinch.	HG	1, 2a	pinch	6.63	2.21
166	Did you hear about the farmer who got attacked by a cow? He milked it for all it was worth.	HG	1, 2a	milk	47.49	1.43
167	Did you hear about the fireman who quit? He said he couldn't take the heat.	HG	1	heat	66.77	1.84
168	Did you hear balloon prices are going up? It's due to inflation.	HG	1	inflation	18.32	1.99
169	Did you hear the joke about the ball? I was rolling around with laughter!	HG	1, 2a, 3	rolling	28.25	4.89
170	Did you know that autopsy is a dying practice?	HG	1	dying	46.27	3.06
171	Digging a hole for a coffin is a grave matter.	HG	1	grave	62.13	3.05
172	Doctors tell us there are over seven million people who are overweight. These, of course, are only round figures.	HG	1, 2a	figure	170.32	3.37
173	Does a shepherd get a staff discount? Or is he just a crook?	HG	1, 2a	crook	5.21	1.73
174	Don't trust people that do acupuncture, they're back stabbers.	HG	1	stab	4.42	3.20
175	Dropped calls are incredibly upsetting, but I'll try not to get hung up on it.	HG	1	hang	37.23	1.88
176	Each time I tried shooting blindfolded I found it an aimless pursuit.	HG	1, 2b, 3	aim	42.54	2.58

177	Early stone tools had many problems that were eventually ironed out.	HG	1	iron	78.55	0.93
178	Every calendar's days are numbered.	HG	1	number	359.35	1.69
179	For every 45 I collected, I would write down the song and the artist so I could keep records.	HG	1, 2b, 3	record	105.77	0.78
180	Frogs have it easy, they can eat what bugs them.	HG	1	bug	10.73	0.08
181	Gardeners like to plant their feet firmly.	HG	1, 2b, 3	plant	57.20	1.04
182	Geology class is the foundation of a decent education.	HG	1, 2a	foundation	56.43	2.03
183	Gravity is studied a lot because it's a very attractive field.	HG	1	attractive	34.98	1.00
184	Having my hair cut for free is the only fringe benefit I receive.	HG	1, 2a, 3	fringe	8.50	0.06
185	He auditioned for a part as a trumpet player but he blew it.	HG	1, 2a, 3	blow	61.64	3.65
186	He couldn't decide whether to accept a job in mattress sales so he decided to sleep on it.	HG	1	sleep	133.50	2.91
187	He had a photographic memory which was never developed.	HG	1, 2a	develop	45.67	2.57
188	He has been a jogger for three years running.	HG	1	run	217.99	2.58
189	I managed to get a good job working for a pool maintenance company, but the work was just too draining.	HG	1, 2a	drain	7.80	0.95
190	I met a girl at an internet cafe, but we didn't click.	HG	1	click	13.94	0.76
191	I never have understood how you're supposed to hit anything with a bat ...it keeps flying away.	HG	1, 2a, 3	bat	10.54	0.04
192	I performed at a prison today. It was a captive audience!	HG	1	captive	8.84	3.72
193	I pitched an idea for making snow tires. It didn't get any traction.	HG	1, 2a	traction	1.64	0.74
194	I probably have blind spots, but I don't see them.	HG	1	see	1404.96	2.80
195	I quit gymnastics because I was tired of hanging around the bars.	HG	1	bar	64.22	0.83
196	I quit gymnastics because the stupid instructor expected me to bend over backwards for her.	HG	1	bend	17.28	1.55
197	I recently spent money on detergent to unclog my kitchen sink. It was money down the drain.	HG	1	drain	7.80	0.95
198	I saw a female deer in my rear-vision mirror. It was case of hind sight.	HG	1	hind	8.42	0.96
199	I see that you have graph paper, you must be plotting something.	HG	1	plot	20.69	1.89

200	I stayed up all night wondering where the sun had gone. Then it dawned on me.	HG	1	dawn	43.15	0.00
201	I thought about becoming a witch, so I tried that for a spell.	HG	1, 2a, 3	spell	29.26	2.94
202	I told the artist that his painting was terrible. I think he got the picture.	HG	1	picture	114.95	2.30
203	I tried wrapping Christmas presents, but I didn't have the gift.	HG	1, 2a	gift	39.22	2.17
204	I used to be a baker, but I didn't make enough dough.	HG	1, 2a	dough	3.58	0.00
205	I used to be a sanitation engineer, but the city dumped me.	HG	1, 2a, 3	dump	4.25	0.67
206	I used to be a tap dancer until I fell in the sink.	HG	1, 2a	tap	15.12	0.76
207	I used to be a transplant surgeon, but my heart just wasn't in it.	HG	1, 2a, 3	heart	369.78	2.96
208	I used to be a Velcro salesman, but couldn't stick with it.	HG	1	stick	52.98	1.99
209	I used to be a watchmaker. It was a great job and I made my own hours.	HG	1, 2a	make	862.00	3.12
210	I used to be addicted to soap, but I'm clean now.	HG	1	clean	72.42	1.85
211	I used to hate math but then I realised decimals have a point.	HG	1, 2a	point	350.84	2.99
212	I used to sell computer parts, but then I lost my drive.	HG	1, 2a, 3	drive	83.67	0.97
213	I used to work at an orange juice factory, but I couldn't concentrate.	HG	1	canned	2.07	2.15
214	I used to work at Starbucks, but I got tired of the daily grind.	HG	1	grind	4.82	0.34
215	I used to work for Budweiser, but then I got canned.	HG	1, 2a	canned	2.07	2.15
216	I used to work in a blanket factory, but it folded.	HG	1, 2a	fold	14.34	0.83
217	I usually take steps to avoid elevators.	HG	1	step	113.28	3.55
218	I went to a buffet dinner with my neighbor, who is a taxidermist. After such a big meal, I was stuffed.	HG	1, 2a	stuff	55.81	1.57
219	A countess started to think about her count-less opportunities.	RH	1	countess	24.09	2.99
220	A day without wordplay is a day without punshine.	RH	1	pun	1.51	0.50
221	A dentist pulled out my tooth without meaning to. It was accidental.	RH	1, 2a	dental	2.86	1.78
222	A Dracula movie had to be reVamped.	RH	1	vampire	1.73	1.61
223	A Freudian slip is when you say one thing but mean your mother.	RH	1, 2b, 3	mother	386.10	2.93
224	A girl entered a hair styling class but failed and didn't make the braid.	RH	1, 2a, 3	braid	2.48	0.71
225	A good pun is its own reword.	RH	1, 2b, 3	reword	0.01	

226	A gossip is someone with a great sense of rumor.	RH	1, 2a, 3	rumor	3.22	0.00
227	A lawyer rips his pants, he goes to a seamstress so she can sue them.	RH	1, 2a	sue	24.98	0.94
228	A lawyer who likes to go fishing is an attorney-at-lure.	RH	1, 2a, 3	lure	4.97	2.48
229	A man was found dead in a vat of falafel condiment. Police are treating it as a hummuscide.	RH	1	homicide	1.50	0.00
230	A medical doctor moonlighted as a theatre critic. When he published a critical review of a production of Madame Butterfly, the director of the show charged him with 'opera rating without a license.'	RH	1	opera	26.14	1.19
231	A paramedic got a new job as a chauffeur: an ambiance driver.	RH	1, 2a, 3	ambience	0.75	0.00
232	A rubber band pistol was confiscated from algebra class because it was a weapon of math disruption.	RH	1, 2b, 3	math	1.01	1.09
233	A thief who stole cutlery without leaving a clue was called the 'stainless stealer'.	RH	1	steal	16.75	2.30
234	Acupuncture is a jab well done.	RH	1, 2a, 3	jab	0.88	1.88
235	After getting pranked by his friends and getting hit with a basket, Aron knew they had a wickered sense of humour.	RH	1, 2a	wicker	2.49	0.01
236	After I used the restroom, I had an out-of-potty experience.	RH	1, 2a	potty	0.87	0.08
237	After the armistice there was a POW exchange. I guess they are right: 'The truce shall set you free.'	RH	1	truce	4.94	0.66
238	After winter, the trees are relieved.	RH	1	relieved	23.54	4.08
239	Alcohol and calculus don't mix so don't drink and derive.	RH	1	derive	23.78	1.21
240	An acquaintance of mine let me try his mixture of basil, olive oil, garlic, and ground pine nuts. We immediately became pesto friends.	RH	1			
241	An authentic diamond is gemuine.	RH	1	gem	4.58	0.00
242	An optometrist told his patient: 'It appears your vision is improving!' 'Really?' replied the patient. 'Must be the luck of the iris.'	RH	1, 2a, 3	iris	4.94	0.02
243	Are you a psychologist? Why are you aFreud to love me?	RH	1	freud	4.75	0.52
244	As a matter of flat, he lives on the 2nd floor.	RH	1	flat	72.39	2.47
245	Asked by a waiter if she would like a drink the lady replied 'wine not?'	RH	1, 2a, 3	wine	76.16	1.00
246	I believe I will be able to run my car on politicians' promises but I'm	RH	1, 2a	fool	59.83	3.11

having trouble with the fool injection system.

247	I met some cult members who worshiped soup serving utensils. I said, 'Oh ye of ladle faith'.	RH	1, 2a, 3	ladle	0.84	1.32
248	I was caught studying the periodic table in English class. It was an elementary mistake.	RH	1, 2a, 3	elemental	3.71	2.41
249	Cell phones are a static symbol.	RH	1, 2b, 3	static	7.03	0.64
250	Cleopatra was the Pharaohs one of all.	RH	1	pharaoh	8.54	1.49
251	I felt sick after I ate the scaloppini. I didn't veal well at all.	RH	1, 2a, 3	veal	1.80	1.23
252	A funeral ship is a sea hearse.	RH	1, 2b, 3	hearse	1.31	0.52
253	A grenade thrown into a kitchen in France would result in Linoleum Blownapart.	RH	1			
254	A new insect extermination company opened last fly day.	RH	1	friday	31.14	1.99
255	An octopus exchanged his old tentacles for new ones. It was 'squid pro quo'.	RH	1, 2a, 3	squid	0.99	1.14
256	Are ubiquitous? No, I ambiguous.	RH	1, 2b, 3	ubiquitous	2.26	1.58
257	Be kind to your dentist. He has fillings, too.	RH	1, 2a, 3	fillings	0.68	1.94
258	Because they moved into an apartment, they didn't need to cut the grass any mow.	RH	1, 2a	mow	1.08	1.29
259	Big spenders have a whole lot of purse-onality.	RH	1	personalty	0.08	
260	Can a physicist read the periodic table? Isotope so.	RH	1	isotope	0.97	0.75
261	Can someone get me a new calendar? Mayan ended.	RH	1	mayan	0.21	
262	Car salesmen compete trying to sell for the lease amount.	RH	1, 2a	lease	12.64	1.85
263	Carpe dime: Seize the ten cent piece.	RH	1	dime	1.73	0.07
264	Coming home tonight I was hit by a semi. I wasn't hurt but I got the freight of my life.	RH	1, 2a	freight	9.27	0.44
265	Cooks who use too many spices could be in-salted.	RH	1	salted	2.03	1.17
266	Couch me if you can, but before that you need to chaise me.	RH	1	couch	16.43	0.00
267	Cows make noise only when they're in the mood.	RH	1, 2a	mood	38.17	3.17
268	Cyclists have lots of wheel-power.	RH	1, 2a	wheel	32.19	1.09
269	Did the electrician want to solve the problem? Yes, some would say it infused him.	RH	1	infuse	0.62	1.49
270	Did you hear about the nervous preacher? He had sweaty psalms.	RH	1	psalm	3.49	0.61
271	Did you hear about the vampire who used to torture his victims with	RH	1	bach		

music? His Bach was worse than his bite.

272	Did you see the movie about the hot dog? It was an Oscar Wiener.	RH	1			
273	Dijon vu -- the same mustard as before.	RH	1			
274	Do you have 11 protons? Because you're sodium fine.	RH	1	sodium	4.40	1.66
275	Drug research companies have created their own cold rush.	RH	1, 2a, 3	cold	169.16	3.19
276	During my trip to Italy, I didn't do much. I just vegged out. When in Rome, do as the Romaines.	RH	1			
277	Even covered in salad dressing my lettuce looked bare, so I put some cloves on it.	RH	1, 2a	cloves	1.42	3.48
278	Ever since Molly moved to Montana she wanted a shiny new kitchen with chrome on the range.	RH	1, 2a	chrome	1.04	1.22
279	Ever since my mineral extraction facility was converted to parking, I've had a lot on my mine.	RH	1, 2a	mine	134.63	3.50
280	Everyone was hungry so we had the pig roaster going full boar.	RH	1, 2a	boar	4.74	2.05
281	Gary said, 'You remind me of a pepper pot.' I said, 'I'll take that as a condiment.'	RH	1, 2a	condiment	0.37	1.51
282	George Bush and Saddam Hussein went to war over Iraq-oncilable differences.	RH	1			
283	Gladys the seamstress was recently inducted into the Pin Pushers Hall of Fame. I guess now she is a status thimble.	RH	1, 2a, 3	thimble	1.21	0.03
284	Good luck to the soccer team! Kick some grass!	RH	1, 2a	grass	73.30	0.06
285	He could play baseball, football, basketball, soccer and tennis. He was a jock of all trades.	RH	1, 2a	word	0.00	0.00
286	He dined with her at the local beanery and was immediately inflatuated.	RH	1	flatulence	0.20	
287	He labored so hard that he worked his fingers to the bonus.	RH	1, 2a	bonus	6.73	0.00
288	He posted an ammunition-for-sale note on the bulletin board.	RH	1	bulletin	6.04	0.20
289	He told me he lost my sieve, but his story didn't hold water.	RH	1, 2b	sieve	1.80	2.27
290	He was upset with his bad start, driving the ball almost beyond the green, but he was able to putt it behind him.	RH	1, 3	put	592.26	2.31
291	He went on a cheese diet in order to cheddar few pounds.	RH	1	cheddar	0.66	1.42
292	Help! I have food stuck in my throat! Ha, just choking.	RH	1, 2b, 3	choking	7.06	2.87

293	How did I know that the small furry animal was not a groundhog? Why, I had to inferret it, of course.	RH	1	infer	5.01	1.98
294	How does a baby get food when it's hungry? Womb service.	RH	1, 2a, 3	womb	5.02	1.05
295	I accidentally left my vacuum cleaner running all night. I guess you could say it was an overnight suck-excess.	RH	1	suck	8.04	2.36
296	I bought my wife a really cute bell ringer for the front door -- it was a-door-a-bell.	RH	1	doorbell	1.37	0.98
297	I can't find the farmer's keys. Maybe I should look in har vest.	RH	1	harvest	15.05	2.23
298	I don't find health-related puns funny anymore since I started suffering from an irony deficiency.	RH	1, 2a, 3	irony	11.01	2.99
299	I don't like cows, they are udderly ridiculous.	RH	1, 2b, 3	udder	0.71	
300	I don't like hanging out at the pancake house that place gives me the crepes!	RH	1, 2a, 3	crepe	0.91	1.08
301	I get distracted by all the meats in the deli section -must be my short attention spam.	RH	1, 2b, 3	spam	1.07	1.14
302	I got a great deal when I bought my apartment. I got the condo minimum.	RH	1	condominium	0.09	
303	I have always wanted to hand out carts at Wal-Mart. I cannot imagine a greeter job.	RH	1	greeter		
304	I just offered someone a job and they accepted, so I offered him my contractulations.	RH	1, 2a, 3	contract	67.84	1.64
305	I keep reading 'The Lord of the Rings' over and over. I guess it's just force of hobbit.	RH	1	hobbit		
306	I knew she was the one when we went on that walk among the evergreen trees. It was love at firs site.	RH	1, 2a, 3	firs	2.12	0.97
307	I knew that masseuse wanted to contact me. She left a massage on my answering machine.	RH	1, 2a, 3	massage	4.09	1.56
308	I like the latest horror movie so much that I've arranged a private screaming.	RH	1, 2a	screaming	13.10	1.53
309	I like to stay current with the electrifying adventures of Sherlock ohms.	RH	1	ohm	0.25	
310	I met the woman of my dreams at the base of Mount Vesuvius. She is the lava my life.	RH	1	lava	7.89	2.38
311	I wanted to be a urologist, but I wasn't good enough. Oh well, I guess urethra got it or you don't.	RH	1			

312	I was enamoured with the famous Paris art museum. It was Louvre at first site	RH	1	louvre	2.46	0.09
313	I was fighting a Dragon, easier slayed then done.	RH	1, 2a	slay	8.47	0.92
314	I was fixated on the pain in my bad tooth. I was abscessed by it.	RH	1, 2a, 3	abscess	0.94	1.95
315	I was kicked out of math class for one too many infractions.	RH	1	fraction	9.40	0.96
316	My neighbor's house was pelted with rotten tomatoes. Police described it as a drive-by fruiting.	RH	1, 2a	fruiting	0.38	0.86
317	The barber opened up a shavings account.	RH	1, 2a, 3			
318	The book about Teflon contained no frictional characters.	RH	1, 2a, 3			
319	A baker stopped making donuts after he got tired of the entire business	CI	1			
320	A backward poet writes in rhymes.	CI	1, 2a			
321	A botanist-turned-prize fighter was penalized for an illegal hit.	CI	1			
322	I finally got rid of that nasty electrical charge I've been carrying. I'm very happy!	CI	1, 2a			
323	A carpenter must have been here. There is a mess.	CI	1			
324	A cardboard belt would be a poor use of paper.	CI	1			
325	A fisherman hated fish and chips but he didn't tell anyone.	CI	1			
326	A farm animal was murdered yesterday. The investigator thinks there was something bad involved.	CI	1			
327	A hawk sat atop a building because it was a bird of prey.	CI	1			
328	A good insurance company knows how to handle claims	CI	1			
329	A man who wanted to sing in church was wondering if he should just ask.	CI	1, 2a			
330	A worker-turned-cook is an assistant chef.	CI	1			
331	A meteor just crashed into Russia ... no comment.	CI	1			
332	A man with no money got into pointless trouble.	CI	1, 2a			
333	After hearing the case of the woman who folded her clothes wrong, the jury had no choice but to sentence her	CI	1, 2a			
334	A young girl in charge of her tribe would be called the little chief.	CI	1			
335	All the waterfowl kept their eyes closed except for one. It was a duck.	CI	1, 2b			

336	After taking the elevator to the top floor I felt very positive.	CI	1, 2a
337	An English teacher, who was dreadfully afraid of insects, while on a picnic screamed like a little girl when he saw there were ants.	CI	1, 2a
338	An electrician knows things which are important.	CI	1
339	Are Philosophy papers graded with scores out of ten?	CI	1
340	Are dog biscuits made of flour?	CI	1, 2a
341	Bakers trade bread recipes on a careful basis.	CI	1
342	Atheism is not a business based organization.	CI	1
343	I felt sick after I ate the scaloppini. I didn't feel well at all.	CI	1
344	Being a baker is hard, you've probably got to take on many different jobs	CI	1
345	Even during a zombie apocalypse, I'd still pick you.	CI	1
346	Bought an apple, bit a piece out of it.	CI	1
347	I bet the butcher $50 that he couldn't reach the meat off the top shelf. He said, 'no, the risk is just too much.'	CI	1
348	I agreed to watch my neighbor's dog, but only if it didn't scratch me. It's in the agreement	CI	1
349	I need to do my philosophy homework but I'm unable to.	CI	1
350	I like to hang my rugs on the clothesline. It is an understandable mistake.	CI	1
351	Did you hear about that new drug that makes people angry? It's very popular right now.	CI	1
352	My supply of pants is being used up.	CI	1, 2a
353	A bad shoemaker's assistant was given his two weeks' notice.	CI	1
354	A baby chicken has a hard time hatching from its egg.	CI	1, 2b
355	A boiled egg in the morning is hard to replace.	CI	1
356	A baseball player can sell himself to a new team if he has a good interview.	CI	1
357	A carpenter sat on his drill and was injured to the point of crying.	CI	1
358	A book called 'Current Trends in Equipping your House' turned out to be a surprising failure.	CI	1

359	A dog gave birth to puppies near the road and was ticketed for making a mess.	CI	1, 2b
360	A doctor who became a bartender was always giving out free drinks!	CI	1
361	A fisherman tried boxing, but he only threw horizontal punches.	CI	1
362	I visit the dentist frequently. So I know what to expect.	CI	1, 2a
363	A worker who moved back to his home town rediscovered his origins.	CI	1
364	A predictable rate is the monthly rent for an apartment.	CI	1
365	A guy walks into a bar with jumper cables. The bartender says, 'You can come in, but don't cause any trouble!'	CI	1
366	Daylight savings is really improving my mood.	CI	1
367	I tried looking for gold, but it didn't work out.	CI	1, 2b
368	A man walks into a bar. The bartender says, 'So, why the long face?'	CI	1, 2a
369	A man was hit by a liquor truck. The drinks were crushing him.	CI	1, 2a
370	A teacher measures lines with a ruler.	CI	1
371	A new batter joined the baseball team, and he was very good at it.	CI	1, 2b
372	I tried wearing tight jeans, but I can never make them work for me	CI	1
373	An egg was late for work, he had to rush to get ready.	CI	1
374	I tried to get a job at a casino but they didn't have a position for me.	CI	1, 2b
375	Aliens can easily understand each other because their language is used everywhere	CI	1
376	Although I like people who are very forward I will never become a surgeon.	CI	1, 2b
377	After 5 years with the same chiropractor, I moved and had to change doctors. It was quite difficult	CI	1
378	After the test drive, the car salesman made his point.	CI	1, 2a
379	After he invented the light bulb, people saw Edison from a new perspective	CI	1, 2a
380	After he ate the duck, the alligator got a little depressed.	CI	1, 2b
381	After my ear operation I feel better.	CI	1, 2a

382	After Junior swallowed the jewelry he had to wait to pass it.	CI	1
383	Artists are colorful people who can take advantage of their emotions.	CI	1
384	After doubts about his vocational calling, the young teacher concluded he was out of his area of expertise	CI	1
385	Cleopatra was the fairest one of all.	CI	1, 2b
386	AT&T and T-Mobile got married, I heard the after party was terrible.	CI	1
387	A day without wordplay is a day without sunshine.	CI	1, 2a
388	A countess started to think about her endless opportunities.	CI	1, 2a
389	A Dracula movie had to be reworked.	CI	1, 2a
390	A dentist pulled out my tooth without meaning to. It was a mistake.	CI	1
391	A girl entered a hair styling class but failed and didn't make the grade.	CI	1
392	A Freudian slip is when you say one thing but mean something else	CI	1
393	A gossip is someone with a great sense of humor.	CI	1
394	A good pun is its own reward	CI	1
395	A lawyer who likes to go fishing is an attorney.	CI	1
396	A lawyer rips his pants, he goes to a seamstress so she can mend them.	CI	1
397	A medical doctor moonlighted as a theatre critic. When he published a critical review of a production of Madame Butterfly, the director of the show charged him with 'working without a license.'	CI	1, 2a
398	A man was found dead in a vat of falafel condiment. Police are treating it as a murder.	CI	1
399	A rubber band pistol was confiscated from algebra class because it was a weapon of disruption.	CI	1
400	A paramedic got a new job as a chauffeur: an ambulance driver.	CI	1
401	Acupuncture is a job well done.	CI	1
402	A thief who stole cutlery without leaving a clue was called the 'cutlery thief'.	CI	1
403	After I used the restroom, I had an unusual experience.	CI	1
404	After getting pranked by his friends and getting hit with a basket, Aron	CI	1

	knew they had a real sense of humour.		
405	After winter, the trees regrow their leaves.	CI	1, 2b
406	After the armistice there was a POW exchange. I guess they are right: 'The truth shall set you free.'	CI	1, 2a
407	An acquaintance of mine let me try his mixture of basil, olive oil, garlic, and ground pine nuts. We immediately became good friends.	CI	1
408	Alcohol and calculus don't mix so don't drink and drive.	CI	1
409	An optometrist told his patient: 'It appears your vision is improving!' 'Really?' replied the patient. 'Must be my good luck'.	CI	1
410	An authentic diamond is real	CI	1, 2a
411	As a matter of fact, he lives on the 2nd floor.	CI	1
412	Are you a psychologist? Why are you afraid to love me?	CI	1
413	I believe I will be able to run my car on politicians' promises but I'm having trouble with the fuel injection system.	CI	1
414	Asked by a waiter if she would like a drink the lady replied 'yes I do'	CI	1
415	I was caught studying the periodic table in English class. It was a basic mistake.	CI	1
416	I met some cult members who worshiped soup serving utensils. I said, 'Oh ye of weak faith.	CI	1
417	Being a poet in prison surely has its advantages and disadvantages.	CI	1
418	Cell phones are a status symbol.	CI	1
419	How did I know that the small furry animal was not a groundhog? Why, I had to deduce it, of course.	CI	2a
420	I met a man who loves eating couches. I think he has a strange tooth.	CI	2a
421	Do you know how winter coats are insulated? They are loaded with feathers.	CI	2a
422	Old actors never die, they just quit their job	CI	2a
423	I changed my iPod's name to titanic and now its working.	CI	2a
424	I love the internet. It's like the friend I never had.	CI	2a
425	Did you know that autopsy is a challenging practice?	CI	2a

426	He couldn't decide whether to accept a job in mattress sales so he decided to pass on it.	CI	2a
427	I performed at a prison today. It was an appreciative audience!	CI	2a
428	I met a girl at an internet cafe, but we didn't get along.	CI	2a
429	I see that you have graph paper, you must be working on something.	CI	2a
430	I used to be a Velcro salesman, but couldn't succeed with it.	CI	2a
431	Did you hear about the fireman who quit? He said he couldn't take the pressure.	CI	2a
432	During my trip to Italy, I didn't do much. I just vegged out. When in Rome, do as the romans do	CI	2a
433	I have always wanted to hand out carts at Wal-Mart. I cannot imagine a preferable job.	CI	2a
434	Did the electrician want to solve the problem? Yes, some would say it excited him.	CI	2a
435	We will continue with the bad jokes	CI	2a
436	I used to work at an orange juice factory, but I was fired	CI	2a
437	I studied a long time to become a doctor, but I didn't have any clients.	CI	2a
438	How do spacemen add more protein to their diet? They make it thicker.	CI	2a
439	A dog not only has a fur coat but also claws	CI	2a
440	I usually take my time to avoid elevators.	CI	2a
441	Big spenders have a whole lot of personality.	CI	2a
442	I wanted to be a urologist, but I wasn't good enough. Oh well, I guess either got it or you don't.	CI	2a
443	Cooks who use too many spices could be insulted.	CI	2a
444	My cousin's girlfriend broke up with him this week. He really cared about her.	LU	2c
445	A friend studied hard but still got a disappointing grade on an exam.	LU	2c
446	Yesterday my alarm didn't go off and I missed an important appointment.	LU	2c
447	My parents recently got a divorce. It has been very challenging to deal with.	LU	2c
448	In the last round of downsizing most of my co-workers were fired. I will probably be next.	LU	2c

449	The car hasn't been running well lately: the other day it wouldn't start at all.	LU	2c
450	I caught a bad flu and spent two weeks in pain and completely bed ridden.	LU	2c
451	I applied for a job I was well qualified for but didn't even get an interview.	LU	2c
452	I worked up my courage to ask out someone I have a crush on but was harshly rejected.	LU	2c
453	My best friend, someone who I really rely on, had to move to the other side of the country.	LU	2c

Appendix B: Example dictionary definitions and latent semantic incongruity estimates.

	Pun	Concept A Definition	Concept B Definition	Incongruity (1-LSA Estimate)
1	Bought an apple, took a byte out of it.	A unit of computer information or data-storage capacity	To press down on or cut into (someone or something) with the teeth	0.8
2	I used to work at an orange juice factory, but I was canned	Preserved in a metal or glass container	To discharge from employment	0.79
3	I used to be a baker, but I didn't make enough dough.	Flour or meal combined with water, milk, etc., in a mass for baking into bread, cake, etc.; paste of bread.	A slang word for money	0.78
4	Did you know that autopsy is a dying practice?	Ceasing to live; approaching death; expiring:	Drawing to a close; ending:	0.76
5	A doctor who became a bartender was always giving out shots!	An injection, as of a vaccine or narcotic drug	A glass of alcoholic drink, especially spirits	0.75
6	A freudian slip is when you say one thing but mean your mother	A female parent	: some other : different from the first or other one	0.75
7	He could play baseball, football, basketball, soccer and tennis. He was a jock of all trades.	A school or college athlete	A person who has many skills : a person who can do many different jobs	0.53
8	I knew that masseuse wanted to contact me. She left a massage on my answering machine.	The action of rubbing or pressing someone's body in a way that helps muscles to relax or reduces pain in muscles and joints	A piece of information that is sent or given to someone	0.53
9	I was kicked out of math class for one too many infractions.	A number (such as 1/2 or 3/4) which indicates that one number is being divided by another	: an act that breaks a rule or law	0.53
10	Have you heard the story about my arm? It's pretty humerus.	The long bone of the upper arm between the shoulder and the elbow	Full of or characterized by humor : funny; causing laughter	0.52
11	How do you organize a space party? You planet.	A large, round object in space (such as the earth) that travels around a star (such as the sun)	A set of actions that have been thought of as a way to do or achieve something ; something that a person intends to do	0.52
12	I used to do rock climbing as a youth, but I was much boulder back then.	A very large stone or rounded piece of rock; a detached and rounded or much-worn mass of rock	Not afraid of danger or difficult situations; showing or needing confidence or lack of fear	0.52
13	Acupuncture is a jab well done.	To push something sharp or hard quickly or suddenly into	The work that a person does regularly in order	0.27

		or toward someone or something	to earn money: a duty, task, or function that someone or something has	
14	I studied a long time to become a doctor, but I didn't have any patience.	The capacity, habit, or fact of being patient; able to remain calm and not become annoyed when waiting for a long time or when dealing with problems or difficult people	An individual or individuals awaiting or under medical care and treatment: the recipient of any of various personal services	0.26
15	He couldn't decide whether to accept a job in mattress sales so he decided to sleep on it.	To take the rest afforded by a suspension of voluntary bodily functions and the natural suspension, complete or partial, of consciousness; cease being awake.	To give (something) extended consideration, especially overnight	0.26
16	Even crazy people know that you should wear hearing protection in high noise areas. That's ear rational.	The characteristic vertebrate organ of hearing and equilibrium; based on facts or reason and not on emotions or feelings	Not thinking clearly : not able to use reason or good judgment	0.25
17	Be true to your teeth, or they will be false to you.	Used as a substitute or supplement, especially temporarily:	Not true or correct; erroneous	0.25
18	Big spenders have a whole lot of purse-onality.	The set of emotional qualities, ways of behaving, etc., that makes a person different from other people	A usually leather or cloth bag used by women for carrying money and personal things : an amount of money that a person, organization, or government has available to use	0.25
19	Each time I tried shooting blindfolded I found it an aimless pursuit.	To position or direct (a firearm, ball, arrow, rocket, etc.) So that, on firing or release, the discharged projectile will hit a target or travel along a certain path.	Having no goal, purpose, or direction	0.24
20	A gossip is someone with a great sense of rumor.	Information or a story that is passed from person to person but has not been proven to be true	A funny or amusing quality : jokes, funny stories, etc., of a particular kind : the ability to be funny or to be amused by things that are funny	0.21
21	I used to be a Velcro salesman, but couldn't stick with it.	To remain attached by adhesion.	To remain persistently or permanently:	0.18

Appendix C: Example of elaboration task provided to participants in study 3.

Screen 1:

"I changed my iPod's name to Titanic and now its syncing."

Please list the words that come to mind when you think of this concept

Syncing: Electronics

IPod, iPhone, merging, joining, iTunes, wires

Screen 2

- "I changed my iPod's name to Titanic and now its syncing."

Please list the words that come to mind when you think of this concept

Sinking: Submerge

Iceberg, death, catastrophe, disaster, cold, Leonardo

Appendix D: Study 1 ethics approval.

Western Research

Western University Health Science Research Ethics Board
NMREB Delegated Initial Approval Notice

Principal Investigator: Prof. Rod Martin
Department & Institution: Social Science\Psychology, Western University

NMREB File Number: 105456
Study Title: The Psycholinguistics of Verbal Puns
Sponsor:

NMREB Initial Approval Date: July 22, 2014
NMREB Expiry Date: April 30, 2015

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Other	Appendix K Survey end questionnaire 6 09 14	2014/06/09
Other	Appendix J. Likert Rating Scale for Pun Study 6 09 14	2014/06/09
Other	Appendix H. pun vs no pun decision 6 09 14	2014/06/09
Other	Appendix E. Item list appendix 6 09 14	2014/06/09
Other	Appendix C. debriefing form 6 09 14	2014/06/09
Western University Protocol	psycholinguistics of verbal puns protocol resubmission clean pdf version 03 07 14	2014/07/03
Recruitment Items	Appendix I. SONA AD clean pdf revision 07 03 14	2014/07/03
Instruments	Appendix G. Demographics questions clean pdf revision 07 03 14	2014/07/03
Recruitment Items	Appendix F. Student Recruitment E-mail for Summer Pool clean pdf revision 07 03 14	2014/07/03
Letter of Information & Consent	Appendix B. gift card version pun study letter of information and consent - clean pdf revision 07 03 14	2014/07/03
Recruitment Items	Appendix D. Recruitment Poster clean pdf revision 07 03 14	2014/07/03
Letter of Information & Consent	Appendix A. gift card version pun study letter of information and consent - clean pdf revision 07 03 14	2014/07/03

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the above named study, as of the HSREB Initial Approval Date noted above.

NMREB approval for this study remains valid until the NMREB Expiry Date noted above, conditional to timely submission and acceptance of HSREB Continuing Ethics Review.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario.

Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00003941.

Ethics Officer, on behalf of Riley Hinson, NMREB Chair

Ethics Officer to Contact for Further Information

Erika Basile ebasile@uwo.ca	Grace Kelly grace.kelly@uwo.ca	Nina Mekhail nmekhail@uwo.ca	Vikki Tran vikki.tran@uwo.ca

This is an official document. Please retain the original in your files.

Appendix E: Study 2 ethics approval.

Western
Research

Western University Health Science Research Ethics Board
NMREB Delegated Initial Approval Notice

Principal Investigator: Dr. Albert Katz
Department & Institution: Social Science\Psychology,Western University

NMREB File Number: 106514
Study Title: Psycholinguistics of Verbal Puns: Thoughts and Feelings
Sponsor:

NMREB Initial Approval Date: April 07, 2015
NMREB Expiry Date: April 07, 2016

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Data Collection Form/Case Report Form	Appendix F. Likert Rating Scale for Pun Study thoughts and feelings	2015/03/01
Data Collection Form/Case Report Form	Appendix G Survey end questionnaire thoughts and feelings	2015/03/01
Other	Appendix E. Item list appendix	2015/03/01
Recruitment Items	Appendix D. SONA AD thoughts and feelings	2015/03/01
Data Collection Form/Case Report Form	Appendix C. Demographics questions thoughts and feelings	2015/03/01
Other	Appendix B. debriefing form thoughts and feelings	2015/03/01
Letter of Information & Consent		2015/03/01
Western University Protocol		2015/03/16

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the above named study, as of the NMREB Initial Approval Date noted above.

NMREB approval for this study remains valid until the NMREB Expiry Date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario.

Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Ethics Officer, on behalf of Riley Hinson, NMREB Chair or delegated board member

Ethics Officer to Contact for Further Information

Erika Basile	Grace Kelly	Mina Mekhail	Vikki Tran
ebasile@uwo.ca	grace.kelly@uwo.ca	mmekhail@uwo.ca	vikki.tran@uwo.ca

This is an official document. Please retain the original in your files.

Appendix F: Study 3 ethics approval.

Western University Health Science Research Ethics Board
NMREB Delegated Initial Approval Notice

Principal Investigator: Dr. Albert Katz
Department & Institution: Social Science\Psychology, Western University

NMREB File Number: 106513
Study Title: Psycholinguistics of Verbal Puns: Elaboration
Sponsor:

NMREB Initial Approval Date: June 02, 2015
NMREB Expiry Date: June 02, 2016

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Data Collection Form/Case Report Form	Appendix F. Elaboration task Likert Rating Scale for Pun Study	2015/03/01
Data Collection Form/Case Report Form	Appendix G Survey end questionnaire-elaboration	2015/03/01
Recruitment Items	Appendix D. SONA AD elaboration	2015/03/01
Other	Appendix E. Item list appendix	2015/03/01
Data Collection Form/Case Report Form	Appendix C. Demographics questions elaboration study	2015/03/01
Other	Appendix B. debriefing form thoughts and feelings	2015/03/01
Revised Letter of Information & Consent		2015/05/22
Revised Western University Protocol		2015/05/22

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the above named study, as of the NMREB Initial Approval Date noted above.

NMREB approval for this study remains valid until the NMREB Expiry Date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario.

Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Ethics Officer, on behalf of Riley Hinson, NMREB Chair or delegated board member

Ethics Officer to Contact for Further Information

___ Eriika Basile	___ Grace Kelly	___ Mina Mekhail	___ Vikki Tran
ebasile@uwo.ca	grace.kelly@uwo.ca	mmekhail@uwo.ca	vikki.tran@uwo.ca

This is an official document. Please retain the original in your files.

Curriculum Vitae

Name: James Boylan

Post-Secondary Education and Degrees:

B.A. (2008) Honours Specialization in Psychology
University of Western Ontario
London, Ontario, Canada

M.Sc. (2011) Cognitive Psychology
University of Western Ontario
London, Ontario, Canada

Ph.D. (2018) Cognitive Psychology
University of Western Ontario
London, Ontario, Canada

Honours and Award:

Western Graduate Research Scholarship
University of Western Ontario
2009-2015

Related Work Experience

Teaching Assistant
The University of Western Ontario
2009-2015

Contract Based Data Analyst
McLean and Company Research Group
London Ontario
2017

Data Analyst
University of Western Ontario
Child Health Research Lab
2018

Publications:

Peer Reviewed

Boylan, J., Katz, A (2013) Ironic Expression can Simultaneously Enhance and Dilute Perception of Criticism. *Discourse Processes*, *50*(3), 187-209.

Frewen, P., Evans, B., Lanius, R., Goodman, J., Halliday, A., Boylan J., Moran, G., Reiss, J., & Lanius, R. (2013) Development of a Childhood Attachment and Relational Trauma Screen: a relational-socioecological framework for surveying attachment security and childhood trauma history. *European Journal of Psychotraumatology.*

Conference Presentations

Boylan, J., Martin, R. and Neufeld, RW. (2015) Latent Semantic Incongruity in Written Puns. Spoken paper presented at 25th Annual Meeting of the Canadian Society for Brain, Behaviour and Cognitive Science, Ottawa, On.

Boylan, J., Martin, R. and Katz, A. (2015) Quantitative Analysis Of Related Text: The Novel Combination Of Latent Semantic Analysis And Multidimensional Scaling. Poster to be presented at the 75th Annual Meeting of the Canadian Psychological Association Ottawa, On.

Boylan, J. (2014). Multinominal process tree modeling approach to the analysis of latent processes in the interpretation of ironic expression of criticism. Symposium talk presented at Western Interdisciplinary student symposium on language research, London On.

Boylan, J., & Katz, A. (2013) Emotion in the organization of autobiographical memory. Poster presented at the 74th Annual Meeting of the Canadian Psychological Association, Quebec City, Que.

Boylan , J., Katz, A. (2012). Ironic criticism simultaneously enhances and dilutes the level of condemnation . Poster presented at the 53 rd Annual Meeting of the Psychonomic Society, Minneapolis, MN.

Boylan , J., Frewen, P. (2012). Analysis of necessary and sufficient symptomology in the diagnosis of PTSD and depression using data from the NCS-R. Poster presented at the 73rd Annual Meeting of the Canadian Psychological Association, Toronto, On.

Boylan, J., & Katz, A. (2012). Analysis of the pragmatic function of sarcasm. Poster presented at the 73rd Annual Meeting of the Canadian Psychological Association, Toronto, On.

Frewen, P., Evans, B., Lanius, R., Moran, G., Goodman, J., Boylan J., & Cloutier, C. (2009). Development of a retrospective computerized attachment & relational trauma scale. Poster presented at CIHR NET Research in Family Violence Meeting, Toronto, On.