

The Effects of User Characteristics on Query Performance in the Presence of Information Request Ambiguity

ABSTRACT

This paper investigates the effects of personality characteristics on individuals' abilities to compose queries from information requests containing various types of ambiguity. In particular, this research examines the effects of user personality characteristics on query performance in the presence of information requests that contained no, extraneous, syntactic, or both extraneous and syntactic ambiguities. The results indicate that personality characteristics significantly affect users' abilities to compose accurate queries. Neuroticism, agreeableness, openness to experience, and conscientiousness significantly affected the number of errors made in the query formulations. Conscientiousness affected the length of time taken to compose the queries and neuroticism affected the confidence users had in the accuracy of their queries. Although several personality dimensions affected query performance, no significant interactions between personality dimensions and ambiguity were detected. Furthermore, both query complexity and information request ambiguity exhibited greater impacts on query performance than personality characteristics. Hence, organizations should attempt to train users to deal with query complexity and information request ambiguity before modifying their training programs for personality characteristics.

(Keywords: Information retrieval, personality, NEO PI-R, ambiguity)

22 **1. INTRODUCTION**

23 In today's highly competitive business environments (Goldstein and Storey, 1994; Cascia and
24 Sanseverino, 1997), organizations are encouraging managers and other end users to query
25 information repositories themselves (Delligatta and Umbaugh, 1993; Owei, 2003). Frequently,
26 the queries these users compose are to satisfy information requests they receive from other
27 stakeholders, e.g., managers, trading partners, and regulatory officials. These information
28 requests are often posed in natural language and typically contain ambiguities such as imprecise
29 adjectives and excessive scope (Whitten et al., 2001). To successfully retrieve the desired
30 information, users require an appropriate skill set (Lerouge et al., 2005). This skill set includes
31 the ability to recognise and ultimately resolve the imprecision contained in an information
32 request. Indeed, good communication skills generally distinguish effective managers from poor
33 managers (Stephens, 1982).

34 Within a business environment, a substantial portion of managerial communication consists of
35 written or verbal requests for information from associates who must transform the request into
36 queries in order to extract data stored in electronic form (Tubre and Collins, 2000). Being in
37 natural language, these information requests often contain ambiguities. Hence, the focus of this
38 research is on the communication ambiguities that may arise between two persons,¹ i.e., the
39 originator of the information (hereafter referred to as the information request provider) and the
40 query developer.

41 Ambiguities in communication, unless resolved, can affect performance (e.g., Cowie and
42 Lehnert, 1996). Furthermore, personality traits can affect the ability of persons to resolve
43 ambiguities (Mumford et al., 1993). Together, these ideas suggest that people with different

¹ Within this research, only verbal ambiguity arising between a information request provided and a query developer is investigated as opposed to situational ambiguity and role ambiguity. Both situational and role ambiguities may also be present within managerial decision making contexts.

44 personality types may cope better with ambiguities and thus achieve better results from database
45 queries than people of other personality types. Accordingly, this paper investigates the effects of
46 personality characteristics on individuals' abilities to resolve ambiguities in an information
47 retrieval task. In particular, this research examines the effects on query performance of the
48 interaction of personality characteristics in the presence of information requests that contained
49 no, extraneous, syntactic, or both extraneous and syntactic ambiguities. The personality
50 dimensions examined are neuroticism, extraversion, openness to experience, agreeableness, and
51 conscientiousness. These dimensions were measured using the Neuroticism, Extraversion,
52 Openness, Personality Inventory-Revised Model (NEO PI-R).

53 Providing insights into the effects of personality characteristics and communication
54 ambiguities on information retrieval is important for a number of reasons. First, the two types of
55 ambiguity investigated here, syntactic and extraneous, occur frequently in organizational
56 communications. These ambiguities exist because of the inability to communicate clearly in
57 conjunction with the tendency not to question our communication skills. Few organizations and
58 individuals adequately recognise that many communications, both internal and external, contain
59 ambiguities. Second, these ambiguities can potentially have disastrous effects on organizational
60 decision making, e.g., ambiguous information requests by stakeholders can lead to incorrect
61 interpretations by the recipients, queries that produce reports that differ from the desired
62 information, and erroneous decisions. By examining whether or not personality can affect an
63 individuals ability to resolve communication ambiguities, organizations can tailor educational
64 programs to reduce the negative effects of these ambiguities. Third, from an academic
65 perspective, little or no research has investigated the potential affects of personality on the ability
66 to resolve ambiguities.

67 The next section of this paper reviews research on personality, ambiguity, and information
68 retrieval, and combines the two areas of research to develop a number of hypotheses
69 investigating the relationships between verbal ambiguity, personality, and information retrieval.
70 Section three explains the research method and section four presents the results. The
71 implications of the research are presented in section five before the conclusions, limitations, and
72 directions for future research are discussed in section six.

73 74 **2. THEORETICAL FOUNDATIONS AND HYPOTHESIS DEVELOPMENT**

75 This section describes the relevance of personality to information systems research and briefly
76 examines one model of personality. The section then discusses the two types of ambiguity
77 investigated within this research and how these ambiguities affect information retrieval. Finally,
78 the section develops five hypotheses that examine the effects of the five personality factors in the
79 presence of syntactic and extraneous ambiguities.

80 81 **2.1 Personality**

82 Personality refers to the cognitive and affective structures maintained by individuals to
83 facilitate their adjustments to the events, people, and situations they encounter (Gough, 1976).
84 Personality variables such as locus of control, ambiguity tolerance, and attitude affect
85 individuals' ability to articulate and evaluate information related tasks. These variables impact
86 perceived usefulness, actual system use, and ultimately MIS success (Zmud, 1979; Taylor,
87 2004). The role of personality on employee behavior has been the focus of organizational
88 psychologists for many years (Johnson, 2003). When carefully applied within the context of
89 specific occupations and organizations, personality variables are significant predictors of job
90 performance (Day and Silverman, 1989; George, 1992; van der Berg and Feij, 2003).

91 Personality variables can potentially increase our understanding of the links between
92 information systems and human cognition (Wheeler et al., 2004). During the 70s and 80s,
93 individual traits, psychological types, and cognitive style were heavily researched within the area
94 of human-computer interaction (Banker and Kauffman, 2004). With the increasing interest in
95 human-computer interaction, Banker and Kauffman (2004) predict a resurgence of this research
96 stream. Recent research into individual traits and computer anxiety (Thatcher and Perrewe,
97 2002) and personal attributes and their impact on adaptation to technological change (Gallivan,
98 2004) support their prediction. When faced with unclear information, personality characteristics
99 impact peoples' ability to provide solutions to problems (Back and Seaker, 2004). Furthermore,
100 personality characteristics have been shown to affect the query errors made during information
101 retrieval (Bowen et al., 2003).

102 A number of models of personality exist. Jung (1962) developed one of the best known
103 theoretical bases used to examine personality and personality types. Jung's theory asserts that
104 individuals have a certain social orientation (either Introverted or Extroverted), prefer one way of
105 perceiving (either Sensing or Intuition), and have one primary way of judging (either Thinking or
106 Feeling) (Johnson et al., 2001). Myers and Briggs added another dimension, during the
107 development of the MBTI, to describe how an individual primarily deals with the outer world
108 (either Judging or Perceiving) (Nordvik, 1996). The resulting Myers Briggs Type Indicator
109 (MBTI) is one widely accepted and reliable measures of personality type (Furnham et al., 2003).

110 Recently another model is receiving considerable attention. The Five Factor Model has
111 been derived from a "more theoretically neutral position" (Widiger and Trull, 1997, pg, 229).
112 While most other models of personality have evolved from one theoretical perspective (as is the
113 case in MBTI), the Five Factor Model has not been derived from one single theory. Rather, the

114 Five Factor Model examined the personality traits that people consider most important in
 115 describing themselves. The predominant instrument for operationalizing the Five Factor Model
 116 is the NEO PI-R (Clark, 2003). This instrument has “demonstrated exceptional psychometric
 117 properties” (Zhang, 2002; pg. 19).

118 The Five Factor model categorizes personality traits into five major dimensions: Neuroticism
 119 (N), Extraversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness
 120 (C). Each dimension is comprised of six facets. The facets contained within each dimension are
 121 described in sections 2.1.1 through 2.1.5 and summarized in Table 1. The creation and
 122 validation of NEO has had a significant impact on research into the effects of personality (Byrne
 123 et al., 2005). The Five-Factor Model of personality has become the dominant basis for
 124 investigating the effects of personality traits (Goldberg, 1993), i.e., the five factor NEO-PI is the
 125 most extensively used measure in recent academic research (Furnham, 1996; Lampe, 2004). The
 126 model has been used in studies of job performance (Barrick and Mount, 1991; Barrick and
 127 Mount, 1993), career success (Judge et al., 1999), and job satisfaction and work adjustment
 128 (Tenopyr, 1993).

129 **Table 1: NEO PI-R dimensions and facets**

Dimension	Neuroticism	Extraversion	Openness to Experience	Agreeableness	Conscientiousness
Facets	Anxiety	Warmth	Fantasy	Trust	Competence
	Hostility	Gregariousness	Aesthetics	Modesty	Self-Discipline
	Depression	Assertiveness	Feelings	Compliance	Achievement-Striving
	Self-consciousness	Activity	Actions	Altruism	Dutifulness
	Impulsive	Excitement-seeking	Ideas	Straightforwardness	Order
	Vulnerability	Positive emotions	Values	Tender-Mindedness	Deliberation

130

131 **2.1.1 Neuroticism**

132 Neuroticism is the tendency to experience negative emotions such as anxiety, hostility, and
 133 depression. From Table 1, the neuroticism dimension is comprised of the following facets:

134 anxiety, hostility, depression, self-consciousness, impulsiveness, and vulnerability. Traits in the
135 neuroticism dimension reflect different ways of reacting to distress circumstances (Costa and
136 McCrae, 1995). Low neuroticism individuals are more able to cope with stressful conditions
137 without becoming overly upset or anxious. Individuals with high levels of neuroticism often
138 have stronger emotional reactions to stressful conditions. Neuroticism has been found to be
139 negatively associated with career success (Judge et al., 1999). The Neuroticism dimension is not
140 included in the MBTI.

141
142 **2.1.2 Extraversion**

143 Extraversion within the Five Factor Model refers to “high activity, sociability, and a tendency
144 to experience positive emotions” (Furnham et al., 2003, pg. 577). From Table 1, the extraversion
145 dimension is comprised of the following facets: warmth, gregariousness, assertiveness, activity,
146 excitement-seeking, and positive emotions. Extraverts are social, active, talkative, upbeat,
147 energetic, and optimistic (Costa and McCrae, 1992). Individuals deficient in these
148 characteristics are termed Introverted. The conceptualization of extraversion embodied in NEO
149 is not the same as the concept of extraversion in Jung’s theory. For example, within NEO,
150 introspection is not related to extraversion but is a characteristic of openness to experience.
151 Extraversion on the five factor model is, however, positively associated with the Extraversion
152 trait and negatively associated with the Introversion trait on the Extraversion-Introversion
153 dimension of MBTI (Furnham et al., 2003). Extraversion, as operationalized by the NEO-PI R,
154 has been shown to be negatively associated with academic performance. Extraversion has also
155 been shown to be related to job performance among certain occupational groups, e.g., salesmen
156 and managers (Hurtz and Donovan, 2000). The performance of other groups, e.g., professionals

157 and skilled workers, does not appear to be strongly associated with the dimension (van der Berg
158 and Feij, 2003).

159
160 **2.1.3 Openness to Experience**

161 Within the Five Factor Model, openness to experience refers to the “tendency to involve in
162 intellectual activities and new experiences” (Furnham et al., 2003, 578). From Table 1, the
163 openness to experience dimension is comprised of the following facets: fantasy, aesthetics,
164 feelings, actions, ideas, and values. The factor relates to intellect, acceptance of novelty, cultural
165 interests, educational aptitude, and creativity, as well as an interest in varied sensory and
166 cognitive experiences (Howard and Howard, 1995). Traits in the openness to experience
167 dimension attempt to capture the process of using cognition, intelligence, and contemplativeness
168 together with unconventionality (Judge et al., 1999). Individuals who score low on this
169 dimension tend to be conventional in behavior and conservative in their outlook, and to prefer
170 familiar and recognisable situations. Low openness to experience individuals also tend to be less
171 creative. Individuals who score high on this dimension are more imaginative and open-minded.
172 They are willing to entertain novel ideas, to think more divergently, and to have positive
173 attitudes towards learning (Costa and McCrae, 1992; Barrick and Mount, 1991). The openness
174 to experience dimension has been shown to be related to training proficiency (Barrick and
175 Mount, 1991). Openness to experience on the five factor model is positively associated with the
176 Intuition trait and negatively associated with the Sensing trait on the Sensing-Intuition dimension
177 of MBTI (Furnham et al., 2003).

178
179 **2.1.4 Agreeableness**

180 Agreeableness within the Five Factor Model refers to “friendly, considerate, and modest
181 behaviour” (Furnham et al., 2003, pg. 578). From Table 1, the agreeableness dimension is
182 comprised of the following facets: trust, modesty, compliance, altruism, straightforwardness,
183 and tender-mindedness. Traits in the agreeableness dimension reflect styles of interpersonal
184 behaviour and interaction (Costa et al., 1991). Individuals with high levels of agreeableness are
185 fundamentally unselfish. They are compassionate and cooperative, and tend to believe others
186 behave in a similar manner (Judge et al., 1999). Individuals with low levels of agreeableness
187 tend to be non-compliant, critical, and sceptical. Agreeableness on the five factor model is
188 positively associated with the Feeling trait and negatively associated with the Thinking trait on
189 the Thinking-Feeling dimension of MBTI (Furnham et al., 2003).

190 **2.1.5 Conscientiousness** 191

192 Conscientiousness within the Five Factor Model refers to “persistence, self-discipline, and the
193 need for achievement” (Furnham et al., 2003, pg. 578). From Table 1, the conscientiousness
194 dimension is comprised of the following facets: competence, self-discipline, achievement-
195 striving, dutifulness, order, and deliberation. Conscientiousness on the five factor model is
196 positively associated with the Judging trait and negatively associated with the Perceiving trait on
197 the Judging-Perceiving dimension of MBTI (Furnham et al., 2003). As such, conscientiousness
198 is a measure of the way persons process information, i.e., whether their Thinking trait or Feeling
199 trait is dominant. Traits in the conscientiousness dimension describe differences in an
200 individual’s motivation and persistence. Individuals with a high level of conscientiousness are
201 competent, productive, and well-organized. As such, they approach problem solving in a highly
202 organised and structured manner that tends to lead to sensible decisions. Individuals with low
203 levels of conscientiousness tend to be prone to procrastination and prefer a “lackadaisical”

204 approach towards achieving their goals. The conscientiousness dimension has been the most
205 consistent predictor of job performance (van der Berg and Feij, 2003).

206
207 **2.2 Ambiguity and Information Retrieval**

208 Research has been conducted into the modelling of data imprecision and data uncertainty
209 within the design of data models and databases (Ma, 2005). Incongruities between information
210 requests and data representations adversely affect end-user accuracy, time taken, and confidence
211 (Borthick et al., 2001). Especially for stakeholders with little prior association with the relevant
212 data repositories, metadata about the entities, relationships, and attributes in these data
213 repositories are often ambiguous. Unfortunately, overcoming the ambiguity/uncertainty
214 associated with the metadata will not necessarily lead to improved decision making, i.e.,
215 resolving metadata ambiguities is a necessary but not sufficient condition for effective
216 information retrieval. Understanding the data model and the associated metadata removes one
217 type of ambiguity, i.e., the query developer must completely understand the data structure with
218 which they are working. The query developer must, however, also understand the information
219 request they are given.

220 Within the realm of information retrieval, a person receives an information request, interprets
221 that request, and formulates a query to retrieve the required information from a data repository,
222 i.e., a database, data mart, or data warehouse. Knowledge workers can access these data
223 repositories via a wide variety of end-user analytical tools including graphical query interfaces,
224 report writers, OLAP cube builders, and data mining tools as well as the more traditional
225 database query languages (Speier and Morris, 2003). The presence of ambiguity in an
226 information request is likely to lead to multiple valid interpretations of the desired information
227 request. Because of the multiplicity of valid interpretations, the information retrieved may not be

228 the information desired by the person making the request. Use of the potentially inappropriate
229 information can have significant negative ramifications on business decision-making processes.

230 Walton (1996) identified six ambiguity types: lexical, syntactical, inflective, pragmatic,
231 emphatic, and suggestive. Axelsen et al. (2001) expanded Walton's taxonomy to include a
232 seventh type of ambiguity, extraneous ambiguity. Axelsen et al. (2001) examined, within one
233 experiment, all seven different types of ambiguity. They found two types of ambiguity, syntactic
234 and extraneous, to significantly affect a person's performance during the query composition
235 process.² Their results indicate that syntactic and extraneous ambiguity strongly affect people's
236 ability to correctly translate information requests into queries that extract the information desired
237 by the requestor. Interestingly, these two types of ambiguities are closely related to Bonner's
238 (1994) classification of task characteristics that are elements of task complexity. Bonner
239 classified task characteristics that are elements of task complexity as relating to either the amount
240 of information (extraneous ambiguity) or clarity of information (syntactic ambiguity). Bonner
241 found both appeared to be negatively related to performance, however, only the effect for clarity
242 of information was significant.

243 Van Gompel et al. (2005) found that globally syntactically ambiguous sentences are as easy to
244 read as sentences containing no ambiguities. They were, however, not investigating the accuracy
245 with which the ambiguities were resolved. Furthermore, they suspected that the person reading
246 the globally ambiguous sentence may be failing to notice or resolve the ambiguity.

247
248 **2.2.1 Syntactic Ambiguity**

² These two types of ambiguity are also considered important because they are two of the most common forms of ambiguity in everyday business communications.

249 Syntactic ambiguity, i.e., structural or grammatical ambiguity, often results in recipients being
250 unclear or mistaken as to the subject or the object of a sentence. An example of syntactical
251 ambiguity occurs in the information request:

252 *“Provide a report of current wine inventory and suppliers that determines their sales*
253 *for the last month.”*

254 The request is syntactically ambiguous, as “their” can refer to either the wine inventory or the
255 suppliers. As illustrated by this example, one of the most common forms of syntactic ambiguity
256 is the use of indefinite pronouns where the pronoun's antecedent is not clear.

257
258 **2.2.2 Extraneous Ambiguity**

259 Extraneous ambiguity arises when information is included that is not required to complete the
260 current task. Some extraneous communications are clearly not relevant to the task at hand, e.g.,
261 small talk, and may even be misleading, e.g., discussions of other projects when one or more
262 participants mistakenly think the discussion does, indeed, affect the current task/project.

263 Axelsen et al. (2001) found that excess information impairs people’s ability to recognise critical
264 elements of an information request. An example of extraneous ambiguity occurs in the
265 information request:

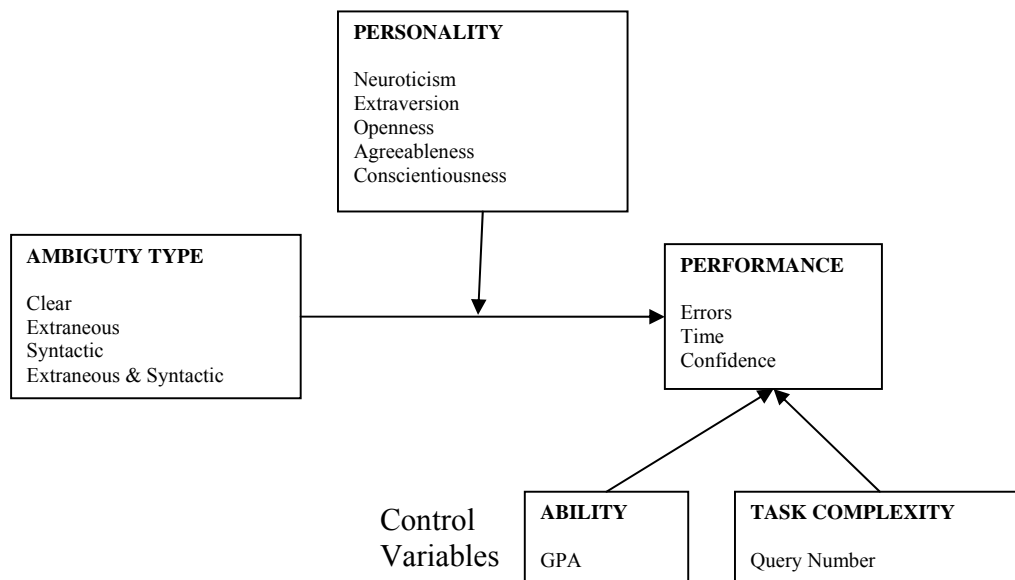
266 *”After the recent spate of increases in premiums, the new insurer wants to know about*
267 *the major risks of theft or damage. Generate a report containing item number, item*
268 *name, item maker, item year, and total value, i.e., quantity on hand times average unit*
269 *cost, where total value is greater \$5,000 or item year is less than 1994.”*

270 The first sentence represents extraneous ambiguity. The sentence is unnecessary but could
271 potentially be beneficial to the recipients by explaining the motivation for the information
272 request. The extraneous information could, however, confuse the recipients and cause them to

273 misinterpret the information request, e.g., by expanding the scope of the query. This research
274 extends the work undertaken by Axelsen et al. (2001) by examining whether some personality
275 types can resolve syntactic and extraneous ambiguities better than other personality types.

276
277 **2.3 Ambiguity, Personality, and Information Retrieval**

278 This section develops a set of testable hypotheses. Each hypothesis examines the effects of
279 the five personality factors on query performance in the presence of syntactic and extraneous
280 ambiguity. The relationships discussed above between ambiguity, personality, and information
281 retrieval performance are depicted in Figure 1.



282 **Figure 1 – Model Tested**

283 **2.3.1 Neuroticism, Ambiguity, and Information Retrieval**

284 Once an information request has been received, to formulate a query users interpret the
285 components of the request relative to the tables and attributes in the data structure. When
286 individuals undertake more demanding attentional tasks, higher levels of neuroticism are
287 associated with worse task performance (Szymura and Wodniecka, 2003; Wallace and Newman,

288 1998). Similar findings from Hertz and Donovan (2003) revealed that lower levels of
289 neuroticism (emotional stability) led to better performance.

290 Introducing ambiguity (either syntactic or extraneous) into an information request increases
291 task complexity (Bonner, 1994) and thus makes the task of formulating the query more
292 demanding and potentially more stressful. Persons with lower levels of neuroticism are better
293 able to remain calm and less anxious when placed in stressful situations (Costa and McCrae,
294 1995). This increased stress associated with ambiguity invokes a negative emotional response in
295 persons with higher levels of neuroticism and negatively affects their performance (Furnham et
296 al., 2002). High neuroticism users faced with both types of ambiguity simultaneously are likely
297 to experience even more difficulty and even greater negative effects on their performance.

298 Within the area of information retrieval three different measures of performance are of
299 interest. First, when placed in more stressful situations, persons with higher levels of
300 neuroticism are less likely to remain calm and thus are more likely to make errors. Second,
301 persons who are able to stay calm in a stressful situation should be able to complete their tasks
302 quicker. Third, persons exhibiting higher levels of anxiety are likely to be less confident in the
303 output of their query. This analysis leads to the following three hypotheses:

304 H1(a): When faced with ambiguity in information requests, users with higher levels of
305 neuroticism will make more semantic errors formulating queries than users with lower
306 levels of neuroticism.

307 H1(b): When faced with ambiguity in information requests, users with higher levels of
308 neuroticism will take longer formulating queries than users with lower levels of
309 neuroticism.

310 H1(c): When faced with ambiguity in information requests, users with higher levels of
311 neuroticism will be less confident with the outcome of their queries than users with lower
312 levels of neuroticism.

313
314 **2.3.2 Extraversion, Ambiguity, and Information Retrieval**

315 Persons with high extraversion scores tend to be more outgoing, high spirited, active,
316 excitement seeking, and cheerful. The relationship between performance and extraversion alters
317 in both significance and direction depending on task and situational variables (Furnham et al.,
318 2002; Hogan and Holland, 2003). The task of composing queries for information requests
319 requires little use of the exuberant traits associated with high levels of extraversion. To perform
320 the task well and to resolve the ambiguity relies on people's ability to focus on concepts and
321 ideas. Research has indicated that introverts have an advantage in written assessments whereas
322 extraverts typically benefit by oral assessment (Furnham et al., 2002). For this research,
323 individuals are required to compose queries for written tasks, thus tending to favor the introverts.
324 Furthermore, the presence of ambiguity within the task is likely to make an extroverts inability to
325 discuss the ambiguity more frustrating. Thus, because they are required to suppress their
326 enthusiasm to engage the external and sensory aspects of a task and to focus intently on the
327 internal and cognitive aspects of the task, individuals exhibiting high levels of extraversion are
328 likely to find the task of composing a query more difficult and stressful,. This increased
329 difficulty is likely to lead to them making more errors and taking more time. Confidence is more
330 problematic. While logic suggests that if a person is experiencing more difficulty, then that
331 person would likely be less confident in the output they produced. Research has shown,
332 however, that higher levels of extraversion lead to overconfidence. Thus, while a direction can

333 be predicted for accuracy and time, a direction for the confidence hypothesis is not possible and
334 thus H2(c) is stated in the null. This analysis leads to the following three hypotheses:

335 H2(a): When faced with ambiguity in information requests, users with higher levels of
336 extraversion will make more semantic errors formulating queries than users with lower
337 levels of extraversion.

338 H2(b): When faced with ambiguity in information requests, users with higher levels of
339 extraversion will take longer formulating queries than users with lower levels of
340 extraversion.

341 H2(c): When faced with ambiguity in information requests, users with higher levels of
342 extraversion will not differ in their confidence in the output of their queries relative to
343 users with lower levels of extraversion.

344
345 **2.3.3 Openness, Ambiguity, and Information Retrieval**

346 Recall that traits in the openness to experience dimension reflect the process of using
347 cognition, intelligence, and contemplativeness together with unconventionality (Judge et al.,
348 1999). Individuals with low levels of openness to experience are more conventional and prefer
349 familiar and recognizable situations. Conversely, individuals with higher levels of openness to
350 experience are likely to have greater ability to achieve innovation, to have more positive attitudes
351 towards learning, and to exhibit higher motivation (Barrick and Mount, 1991). These individuals
352 are also more willing to embrace novel ideas “as well as experience emotions more keenly”
353 (Howard and Howard, 1995, pg. 15). Prior research into cognitive style factors that affect
354 database query performance find that persons who rely on intuition as opposed to sensing

355 composed more accurate queries (Bowen et al., 2003).³ As noted previously, a positive
356 relationship exists between the openness dimension in the Five Factor Model and the Intuition
357 trait within the MBTI (Furnham et al., 2002).

358 Determining and extracting the information required from an information system requires
359 creative mappings of real world ideas and concepts to a database structure (Wand and Weber,
360 1990). Based on prior research, individuals with higher levels of openness are more likely to
361 compose more accurate queries (Furnham et al., 2002; Bowen et al., 2003). Individuals with
362 higher levels of openness to experience exhibit higher levels of flexibility and creativity and,
363 hence, should find the task less daunting. When higher openness individuals are also confronted
364 with excess information and structurally unsound grammar (extraneous and syntactic ambiguity),
365 their greater flexibility and creativity increase the likelihood that such individuals can resolve the
366 ambiguity and perform better than individuals who are less open.⁴ That is, individuals who score
367 higher on the openness dimension are likely to be able to produce more accurate queries in less
368 time and be more confident in their output. This analysis leads to the following three
369 hypotheses:

370 H3(a): When faced with ambiguity in information requests, users with higher levels of openness
371 to experience will make fewer semantic errors formulating queries than users with lower
372 levels of openness to experience.

³ Bowen et al., (2003) used the MBTI instrument to measure the dimensions of personality.

⁴ Recent research into the area of personality research has identified that a possible reason for poor relationship between openness to experience and job performance may be due to the presence of two subfactors affecting performance differently (Griffin and Hesketh, 2004). Openness-to-experience may be comprised of one factor that relates to openness to internal experience and one factor that relates to openness to external experience. Thus, the openness to external factors is likely to be related to the adaptability required for the task of querying. As such, the current research has couched the hypotheses in terms of the long standing five factor model, but, at the same time, when dealing with the interpretation of the results the possibility of the sixth factor will be taken into consideration.

373 H3(b): When faced with ambiguity in information requests, users with higher levels of openness
374 to experience will take less time formulating queries than users with lower levels of
375 openness to experience.

376 H3(c): When faced with ambiguity in information requests, users with higher levels of openness
377 to experience will be more confident with their query output than users with lower levels
378 of openness to experience.

379
380 **2.3.4 Agreeableness, Ambiguity, and Information Retrieval**

381 Recall that individuals with high levels of agreeableness are compassionate and cooperative
382 whereas individuals with low levels of agreeableness tend to be more non-compliant, critical,
383 sceptical, and competitive (Costa et al., 1991; Judge et al., 1999). The process of query
384 composition requires that essential information in the information request is recognised and that
385 individuals step through the components of each query logically and cautiously. Especially due
386 to their propensity to be critical, sceptical, and competitive, individuals exhibiting lower levels of
387 agreeableness are likely to be better able to recognise, articulate, and evaluate the information
388 necessary to make accurate analyses. Conversely, agreeable individuals with higher levels of
389 straightforwardness, ingenuousness, and modesty are more likely to misinterpret and overlook
390 relevant information (Costa et al., 1991; Judge et al., 1999).

391 Following the execution of each query, users must evaluate, on an objective and logical basis,
392 the accuracy and relevance of the results generated. The difficulty of these tasks increases with
393 excess information or structural ambiguities. Individuals who are less agreeable, exhibit greater
394 scepticism, and exercise critical thinking skills are more likely to recognise the presence of
395 ambiguities and to be better equipped to resolve them. If the participant's personality type is
396 better suited to the task, they are likely to exhibit better performance by way of accuracy, take

397 less time, and be more confidence in their output.⁵ This analysis leads to the following three
398 hypotheses:

399 H4(a): When faced with ambiguity in information requests, users with higher levels of
400 agreeableness make more semantic errors formulating queries than users with lower
401 levels of agreeableness.

402 H4(b): When faced with ambiguity in information requests, users with higher levels of
403 agreeableness take more time formulating queries than users with lower levels of
404 agreeableness.

405 H4(c): When faced with ambiguity in information requests, users with higher levels of
406 agreeableness are less confident in their query output than users with lower levels of
407 agreeableness.

408

409 **2.3.5 Conscientiousness, Ambiguity, and Information Retrieval**

410 The process of composing queries from information requests is iterative. Individuals with
411 higher levels of conscientiousness possess “persistent and achievement oriented” traits (Bryne et
412 al., 2005). When presented with excess information or syntactical ambiguity in an information
413 request, users with higher levels of conscientiousness are more likely to carefully, logically, and
414 persistently work through the request. Furthermore, because they are achievement-oriented and
415 thus more diligent, they are likely to produce more accurate queries and to be more confident in
416 their query results. Their greater diligence is, however, likely to increase the amount of time
417 they spend formulating their queries. This discussion leads to the following hypotheses:

⁵ Recent research into the area of personality research has identified the possibility of a sixth factor referred to as the honesty-humility factor (Ashton and Lee, 2005). This factor is derived from two of the facets within the agreeableness dimension. Due to the nature of the experimental task forming part of a participant’s assessment the affects of this possible dimension should be minimized. As such, the current research has couched the hypotheses in terms of the long standing five factor model, but at the same time when dealing with the interpretation of the results the possibility of the sixth factor will be taken into consideration.

418 H5(a): When faced with ambiguity in information requests, users with higher levels of
419 conscientiousness make fewer semantic errors formulating database queries than users with
420 lower levels of conscientiousness.

421 H5(b): When faced with ambiguity in information requests, users with higher levels of
422 conscientiousness take more time formulating database queries than users with lower levels
423 of conscientiousness.

424 H5(c): When faced with ambiguity in information requests, users with higher levels of
425 conscientiousness are more confident in their query output than users with lower levels of
426 conscientiousness.

427

428 **3. METHOD**

429 **3.1 Research Design, Participants, and Data Collection**

430 In a laboratory experiment, participants composed and executed queries in SQL for an Oracle
431 database⁶. Seventy-five undergraduate and masters level commerce students participated in the
432 experiment. All participants were familiar with general computing concepts and activities and,
433 prior to the experiment, had received training in developing SQL queries. All participants
434 received a set of instructions containing the scenario, the details of tasks to be performed, the
435 data dictionary, and the entity-relationship diagram (Appendix A). To eliminate potentially
436 different interpretations of non-verbal cues that accompany face-to-face verbal communication
437 (Manusov et al., 1997), ambiguity was manipulated via written information requests. The
438 objective of this research is to study the impacts of personality variables on resolving syntactic,
439 extraneous, and the combination of syntactic and extraneous ambiguity on information retrieval

⁶ While many users within today's environment use applications that incorporate a querying by example (QBE) tool the authors have followed the advice by Hayes and Hunton, (2001) who state that "although QBE tools are visual and relatively easy to use, they're somewhat limited. To create complex queries users must turn to a language call SQL." They go on to say that "in today's world it is important to understand the fundamentals of SQL, for it is the basis of all database queries."

440 performance.⁷ Each information request was designed with four formulations, one clear (no
 441 ambiguity) and the remaining three formulations corresponding to each type of ambiguity.
 442 Figure 2 provides an example of an information request, in each of the four formulations, and the
 443 corresponding model query. Participants received information requests in each of four possible
 444 states: clear, extraneous, syntactic, and both extraneous and clear, i.e., each participant
 445 experienced each type of ambiguity.

446 Four equivalent groups were established using the following technique. The participants were
 447 ranked in descending order according to their GPA,⁸ i.e., the person with the highest GPA was
 448 ranked 1 and the next ranked 2, etc. Participants were assigned to four groups according to their
 449 rank, i.e., the highest ranked person to group 1, the second highest to 2, third to group 3, fourth to
 450 group 4, fifth to group 4, sixth to group 3, etc. This method of randomization was intended to
 451 make the overall ability of the groups as equivalent as possible. The groups were then randomly
 452 assigned to a different starting treatment. Thus, participants in each of the four groups
 453 experienced each type of ambiguity in the same order (i.e., order was deliberately fixed) but with
 454 different starting points. Table 2 shows the order in which the ambiguities were presented to the
 455 participants in each group.⁹

456 **Table 2: Information request matrix illustrating type of ambiguity present within the information request**

Information Request	Group A N=18	Group B N=20	Group C N=21	Group D N=16
1	Clear	Extraneous	Syntactic	Both
2	Both	Clear	Extraneous	Syntactic
3	Syntactic	Both	Clear	Extraneous
4	Extraneous	Syntactic	Both	Clear
5	Clear	Extraneous	Syntactic	Both
6	Both	Clear	Extraneous	Syntactic

⁷ Recall information retrieval performance is defined within this study as (1) the accuracy (effectiveness) of query formulation, (2) the time required to formulate queries, and (3) users' confidence in their query formulations.

⁸ The GPA for each participant was obtained from the university, i.e., not self reported. Due to the possibility that some students were in their first semester of study at the current university, the preferred measure of grade point average for IS/IT courses was not obtainable for all students. Overall GPA was used as the best available alternative.

⁹ Group was not significantly associated with the number of errors made or the time taken.

7	Syntactic	Both	Clear	Extraneous
8	Extraneous	Syntactic	Both	Clear
9	Clear	Extraneous	Syntactic	Both
10	Both	Clear	Extraneous	Syntactic
11	Syntactic	Both	Clear	Extraneous
12	Extraneous	Syntactic	Both	Clear

457

Information Request 2	
Formulation	Information Request
Clear	List item number, item name, quantity on hand, and quantity on order for those items where the quantity on hand is greater than 2 times the quantity ordered.
Syntactic	Management wants a list of inventory items, names, and quantities where the stock levels and the typical amounts ordered are double. <i>(The ambiguity is caused by the use of “and” in the phrase “where the stock levels and the typical amounts ordered are double”. It is unclear as to which amount is to be doubled – the amounts ordered or the stock levels.)</i>
Extraneous	A recent stocktake of a random sample of inventory items revealed some significant shortages. Provide management with a list containing item number, item name, quantity on hand, and quantity on order for those items where the quantity on hand is greater than 2 times the quantity ordered.
Syntactic and Extraneous	A recent stocktake of a random sample of inventory items revealed some significant discrepancies. Provide management with a report of inventory items, names, and quantities where the stock levels and the typical amounts ordered are double.
<i>Model Answer</i> SELECT item_no, item_name, qty_hand, qty_ordered FROM inventory WHERE qty_hand > 2 * qty_ordered;	

458 **Figure 2: Example of an information request and the four different formulations**

459

460 The participants had two hours to construct, as accurately as possible, appropriate queries for as
 461 many of the twelve information requests as they could (Appendix B). Participants received 5%
 462 course credit for participating. Participants were informed that they would be marked on the
 463 accuracy of the queries they entered and not merely the number of queries they completed.
 464 Because the correct query formulations were generally increasing in complexity, participants
 465 were encouraged to do their best on each query before moving to the next information request¹⁰.
 466 Participants used a UNIX shell script that recorded their entire session. After submitting each

¹⁰ The grading criteria for the students' results, not the coding for the statistical analysis, were as follows. The students received a base of 50% of the available 100 points if they produced at least four syntactically correct queries that reasonably addressed the corresponding information requests. Essentially all students received this 50 points. Each completed query was graded on a 0 to 5 scale based on its accuracy. Because of the increasing complexity of the queries, obtaining the same score on each successive query became increasingly challenging.

467 query attempt, the system displayed the SQL result, i.e., either the rows returned by the query or
468 a syntax error message. Participants could revise their queries as many times as they wished.
469 When they indicated that they were satisfied with the result they obtained for a particular request,
470 participants were prompted to specify their confidence that the query results were correct. After
471 indicating their confidence level, participants proceeded to the next information request.

472 Note that within this study, the ambiguities contained within the information request could
473 have been resolved by the query developer through the use of the additional information
474 provided e.g., the data dictionary. As such, even though there were ambiguities, e.g., semantic
475 ambiguity, within an information request there is only one possible semantically correct
476 interpretation in the context of this experiment.

477 478 **3.2 Operationalizing the Variables**

479 The dependent variables were, for each query developed by each participant: the number of
480 semantic errors in the query, the time taken to compose the query, and the participant's
481 confidence in their query. The number of semantic errors was determined by counting the
482 number of semantic errors in each participant's last query attempt for each information request.
483 Information requests that were not attempted by a particular participant were not included in the
484 scoring. Furthermore, the final question being attempted at the end of the two hour period may
485 not have been included when it was obvious that the SQL query was not complete. After two
486 individuals independently counted query errors, they cross-checked their error coding sheets for
487 correctness and consistency. When the two coders compared their solutions, the possible
488 outcomes were initial total agreement, one coder being deemed correct, or both coders changing
489 their solution. Given the criteria of making the minimum number changes to reach a
490 semantically correct solution, after re-examining each query and each coder's solution, the coders

491 were always able to reach agreement on the number of errors, if any, in each participant's
492 query¹¹.

493 The time taken to compose the query for the information request was determined by
494 examining the log files. The dependent variable for the participant's self-assessed confidence
495 level was entered on the scale: 86-100%, 71-85%, 56-70%, 41-55%, 26-40%, 11-25%, and 0-
496 10%. Participants were asked to determine which of the seven categories they considered best
497 represented their level of confidence. The values were transformed to a seven point scale as
498 follows: ratings of 86-100% transformed to 7, ratings of 71-85% to 6, ratings of 56-70% to 5,
499 ratings of 41-55% to 4, ratings of 26-40% to 3, ratings of 11-25% to 2, and ratings of 0-10% to 1.

500

501 **3.2.1 Independent Variables**

502 The six primary independent variables were ambiguity type and each participant's scores on
503 each of the five dimensions of the NEO PI-R. Ambiguity type was coded as a categorical
504 variable taking on one of four values: clear, extraneous, syntactic, or both extraneous and
505 syntactic.

506 Prior to the experiment all participants completed a Neuroticism, Extraversion, Openness,
507 Personality, Inventory-Revised Model (NEO PI-R) survey to determine their personality types.
508 Each participant completed a self-reported item booklet (Form S) using the hand scoring answer
509 sheet. Due to copyright restrictions, copies of the Form S (completed by participants), the item
510 booklet, and summary sheets are not included in the Appendices.¹² Each participant's score on

¹¹ To give an indication of the process, the two coders recorded the outcomes for 15% of the queries. The coders were in initial total agreement on 85% of the queries. On the remaining 15%, the more experienced coder was deemed correct on 50%, the less experienced coder was deemed correct on 33.3% of the queries, and both coders changed their error counts on 16.7% of the queries.

¹² The documents used were obtained via the Australian Council for Educational Resources see <http://shop.acer.edu.au/acer-shop/locate?group=RQ>

511 each of the five dimensions of the NEO PI-R was calculated by hand using the procedures
512 outlined in the manual (Costa and McCrae, 1995).

513 To control for task complexity and each user's query ability, two more variables, query
514 number and grade point average, were used as covariates in the statistical analyzes. Because the
515 model queries that satisfied the 12 information requests became increasingly more challenging,
516 query number was selected as the proxy for query complexity. Alternate complexity measures
517 such as length, difficulty, and effort metrics (Halstead, 1977) were not chosen because they focus
518 almost exclusively on size. The order of the information requests took into account the
519 "challenges" encountered by participants when composing a query. A query containing a sub
520 query or outer join, for example, is likely to be shorter than a query joining multiple tables.
521 Participants, however, often find the shorter query more difficult and thus more challenging.
522 Because of its greater availability, consistency, comparability, and verifiability, grade point
523 average (GPA), instead of number of IS/IT subjects taken by each participant, was chosen as the
524 proxy for query ability. GPA was obtained from university records.

525 526 **4. RESULTS**

527 **4.1 Summary Results**

528 Table 3, Panel A summarizes the participants' characteristics. Table 3, Panel B summarizes the
529 performance for each group. Table 3, Panel C summarizes, by ambiguity type, the performance
530 of the participants. These results indicate that query developers' effectiveness and confidence
531 were affected primarily by syntactic ambiguity but not by extraneous ambiguity. The means for
532 the number of semantic errors and confidence for the clear and extraneous categories are not
533 statistically different (semantic errors $p = 0.9992$; confidence $p = 0.9753$). Similarly, the mean
534 number of semantic errors and confidence for the syntactic and both categories are not

535 statistically different (semantic errors $p = 0.2555$; confidence $p = 0.6486$). Thus, the variation
536 in the number of semantic errors and confidence appears to stem primarily from syntactic
537 ambiguity. The mean number of semantic errors and confidence for the extraneous and syntactic
538 categories are statistically different (semantic errors $p = 0,0018$; confidence $p = 0.0540$).

539 The results indicate that syntactic ambiguity alone had only a marginal affect on time. The
540 means for time taken for the clear and extraneous categories are not statistically different ($p =$
541 0.3293). The means for time taken for the extraneous and syntactic categories are also not
542 statistically different ($p = 0.2797$). When, however, information requests contained both
543 syntactic and extraneous ambiguity, query developers took significantly longer than when faced
544 with only one type of ambiguity. The means for time taken for the syntactic and both categories
545 are statistically different ($p = 0.0096$). One approach people often take to reduce ambiguity in
546 an information request is by including additional explanatory information. This research shows
547 that if the initial request also contained syntactic ambiguity, the inclusion of extra information
548 was not helpful relative to resolving the syntactic ambiguity, i.e., query developers did not make
549 significantly fewer errors.

550 **Table 3 – Summary Results**

551

Panel A - Participant Characteristics

Group A	Group B	Group C	Group D
----------------	----------------	----------------	----------------

Grade Point Average (7-point scale, 7 highest)				
Mean	4.94	5.08	4.90	4.86
Standard deviation	0.69	0.74	0.74	0.79
Gender				
Number of males	11	9	11	12
Number of females	7	11	10	4
Neuroticism				
Mean	52.22	54.70	51.19	57.50
Standard deviation	9.53	12.52	9.41	13.98
Extroversion				
Mean	51.78	53.20	51.86	47.07
Standard deviation	7.96	10.53	10.80	11.23
Openness to experience				
Mean	50.17	55.10	52.14	51.25
Standard deviation	8.54	9.51	9.74	10.21
Agreeableness				
Mean	46.28	49.60	46.24	47.50
Standard deviation	9.50	7.86	7.48	11.22
Conscientiousness				
Mean	44.11	44.70	44.00	49.00
Standard deviation	9.02	13.49	10.07	7.60

552

553

Panel B - Participants Summary Performance¹³

	Group A	Group B	Group C	Group D
Semantic Errors/Query				
Mean	3.7153	3.4860	3.5758	4.5952
Standard deviation	5.7630	5.9647	6.4412	6.409
Time Taken/Query				
Mean	12.6761	11.4131	12.6934	12.6098
Standard deviation	7.8827	8.1831	7.7507	6.2585
Confidence/Query				
Mean	5.8125	5.4693	6.0969	5.7143
Standard deviation	1.3891	1.7297	1.2698	1.4303

554

555

556

Panel C - Ambiguity type and Summary Performance

	Clear	Extraneous	Syntactic	Both
Semantic Errors/Query				
Mean	2.5263	2.5256	4.6753	5.4605
Standard deviation	5.2005	5.7434	5.9441	7.0962
Time Taken/Query				
Mean	10.6579	11.4934	12.4165	14.6475
Standard deviation	7.6038	6.8310	6.9677	8.5389
Confidence/Query				
Mean	5.9539	5.9487	5.6234	5.5461
Standard deviation	1.4111	1.2791	1.6571	1.5604

557

558

559

560

561

¹ When group is included as an independent variable, the results of the MANCOVA do not reveal a significant association between group number and the number of errors made and time taken.

¹³ When group is included as an independent variable, the results of the MANCOVA do not reveal a significant association between group number and the number of errors made and time taken.

562 **Effects of Personality and Ambiguity of Semantic Errors Made By Users During Query**

563 **Composition**

564 Hypotheses 1(a) through 5(a) predicted that persons possessing various personality
565 characteristics would be more or less successful in formulating queries for information requests
566 containing no, syntactic, extraneous, or both syntactic and extraneous ambiguities. None of the
567 interactions between the five personality dimensions and the four types of query formulations
568 (clear and the three ambiguous) are significant. That is, individuals with various levels of the
569 five different personality dimensions were not significantly more or less successful in resolving
570 ambiguities contained within the information requests. The results of a multivariate analysis of
571 covariance (MANCOVA) (Table 4, Panel A) indicate, however, significant associations between
572 four of the five personality dimensions and number of query errors made. In particular, the
573 results indicate that neuroticism ($F_{1,603}=4.11$, $p=0.0430$, two-tail test), openness ($F_{1,603}=4.75$,
574 $p=0.0297$, two-tail test), and agreeableness ($F_{1,603}=5.23$, $p=0.0226$, two-tail test) significantly
575 affected the number of semantic errors. Conscientiousness has a marginal affect on the number
576 of semantic errors ($F_{1,603}=2.87$, $p=0.0906$, two-tail test).

577 The parameter estimates for Openness and Agreeableness are in the directions predicted, i.e.,
578 persons who exhibited higher levels of openness made fewer errors and persons who were more
579 agreeable made more errors. The directions of the parameter estimates for neuroticism and
580 conscientiousness are opposite to those predicted. For neuroticism, the sign of the parameter
581 estimate indicates that persons with lower levels of neuroticism made more errors. Recall that
582 persons who score low on neuroticism are “calm, even-tempered, and relaxed, and are able to
583 face stressful situations without becoming upset or rattled” (Costa and McCrae, 1992, pg. 15).
584 Low neuroticism individuals in this study made fewer attempts to compose queries and made

585 more errors. This outcome suggests that the relaxed attitude of lower neuroticism individuals
 586 tended to make them assume their queries were correct sooner than higher neuroticism
 587 individuals, e.g., as soon as they eliminated all syntax errors.

588 For conscientiousness, the direction of the parameter estimates indicates that, *ceteris paribus*,
 589 persons with higher levels of conscientiousness made more errors. One possible explanation of
 590 this finding is the use of GPA as a covariate. For example, if there are two individuals with the
 591 same level of conscientiousness and one has a higher intellect, then the individual with the higher
 592 intellect should attain a higher GPA. Equivalently, if two students have the same GPA but
 593 different levels of conscientiousness, then the student with the lower level of conscientiousness is
 594 likely to possess the higher intellect¹⁴.

595 Post hoc analysis was performed to examine the contribution of each variable toward the
 596 number of semantic errors. Seven of the eight independent variables were significant. Of these
 597 seven significant variables, complexity and ambiguity, together explained approximately 11.4%
 598 of the variance in performance. The five personality variables together explained only
 599 approximately 2% of the variance in performance. Thus, while some personality variables did
 600 significantly affect the number of errors made, their practical importance is unclear.

601 **Table 4: Effect of Types of Personality and Ambiguity on Performance**

602

Panel A - Effect of Model on Number of Semantic Errors Made During Query Composition						
Source	R ²	df	Mean Square	F Value	Pr > F	Parameter Estimate
Model	0.1501	10	348.99	10.65	0.0001	
Error		603	32.76			
AMBIGUITY		3	338.72	10.34	0.0001	
NEUROTICISM		1	134.75	4.11	0.0430	-0.0417
EXTRAVERSION		1	9.05	0.25	0.6203	-0.0116
OPENNESS		1	155.52	4.75	0.0297	-0.0545
AGREEABLENESS		1	171.17	5.23	0.0226	0.0626
CONSCIENTIOUSNESS		1	94.10	2.87	0.0906	0.0399
QUERY NUMBER ⁺		1	1957.83	59.77	0.0001	0.6664

¹⁴ Formally, assume that conscientiousness * intelligence = GPA and that two students have the same GPA, i.e., conscientiousness1 * intelligence1 = conscientiousness2 * intelligence2 = GPA. For these two students with the same GPA, if student 1 is more conscientious, i.e., if conscientiousness1 > conscientiousness2, then intelligence1 < intelligence2, i.e., student 1 is less intelligent than student 2 or, equivalently, the less conscientious student is more intelligent than the more conscientious student.

603
604

GPA [@]	1	165.18	5.04	0.0251	-0.7710
------------------	---	--------	------	--------	---------

Panel B - Effect of Model on Number of Time Taken During Query Composition

Source	R ²	df	Mean Square	F Value	Pr > F	Parameter Estimate
Model	0.0757	10	270.57	4.49	0.0001	
Error		603	54.80			
AMBIGUITY		3	420.76	7.68	0.0001	
NEUROTICISM		1	5.89	0.11	0.7432	0.0087
EXTRAVERSION		1	10.98	0.20	0.6545	-0.0135
OPENNESS		1	52.66	0.96	0.3274	-0.0317
AGREEABLENESS		1	0.00	0.00	0.9951	0.0002
CONSCIENTIOUSNESS		1	292.53	5.34	0.0212	0.0703
QUERY NUMBER ⁺		1	200.84	3.66	0.0560	0.2134
GPA [@]		1	990.63	18.08	0.0001	-1.8882

605
606

Panel C - Effect of Model User Confidence During Query Composition

Source	R ²	df	Mean Square	F Value	Pr > F	Parameter Estimate
Model	0.0805	10	10.98	5.28	0.0001	
Error		603	2.08			
AMBIGUITY		3	6.60	3.17	0.0238	
NEUROTICISM		1	18.86	9.07	0.0027	-0.0156
EXTRAVERSION		1	0.44	0.21	0.6470	-0.0027
OPENNESS		1	0.43	0.21	0.6496	0.0029
AGREEABLENESS		1	0.63	0.30	0.5825	-0.0038
CONSCIENTIOUSNESS		1	1.25	0.60	0.4387	0.0046
QUERY NUMBER ⁺		1	63.53	30.56	0.0001	-0.1200
GPA [@]		1	5.76	2.77	0.0965	0.1440

607
608
609

⁺ query number is a covariate to proxy for complexity and is the number of the query i.e., 1 to 12

[@] GPA is the students overall GPA and is a covariate to proxy for intellect

610

4.2 Effects of Personality and Ambiguity on Time During Query Composition

611

Hypotheses 1(b) through 5(b) predicted that persons possessing various personality

612

characteristics would be more or less efficient in formulating queries for information requests

613

containing no, syntactic, extraneous, or both syntactic and extraneous ambiguities. None of the

614

interactions between the five personality dimensions and the four types of query formulations

615

(clear and the three ambiguous) were significant. That is, individuals with various levels of the

616

five different personality dimensions did not take significantly more or less time to construct the

617

queries from information requests containing different types of ambiguity. The results of the

618

MANCOVA, reported in Table 4 Panel B, indicate, however, a significant association between

619

one of the five personality dimensions and the time taken to construct the queries. In particular,

620

the results indicate that conscientiousness had a significant effect on the time taken to compose

621 queries ($F_{1,603}=5.34$, $p=0.0212$, two-tail test). The parameter estimates show that, as predicted,
622 persons exhibiting higher levels of conscientiousness took longer to complete each query.

623 Post hoc analysis was performed to examine the contribution of each variable toward time
624 taken to formulate each query. Four of the eight independent variables, with only one variable
625 related to personality, were significant. Of these four significant variables, the three not related to
626 personality contributed the majority of the R^2 , i.e., approximately 7%. Thus, while one
627 personality variable, conscientiousness, did significantly affect time taken, its overall
628 contribution to explaining variations in time was minimal (less than 1%).

629
630 **4.3 Effects of Personality and Ambiguity on User Confidence During Query**
631 **Composition**

632 Hypotheses 1(c) through 5(c) predicted that persons possessing various personality
633 characteristics would exhibit different levels of confidence in the accuracy of their queries for
634 information requests containing no, syntactic, extraneous, or both syntactic and extraneous
635 ambiguities. None of the interactions between the five personality dimensions and the four types
636 of query formulations (clear and the three ambiguous) were significant. That is, individuals with
637 various levels of the five different personality dimensions were not significantly more or less
638 confident in the queries they produced from information requests containing various types of
639 ambiguity. The results of the MANCOVA, reported in Table 4 Panel C, indicate, however, a
640 significant association between one of the five personality dimensions (neuroticism) and
641 confidence. In particular, the results indicate that neuroticism ($F_{1,603}=9.07$, $p=0.0027$, two-tail
642 test) significantly affected the confidence that users had in the accuracy of their queries. The
643 parameter estimate shows that, as predicted, persons with higher levels of neuroticism were less
644 confident in the accuracy of their queries.

645 Post hoc analysis was performed to examine the contribution of each variable toward the
 646 confidence of the users. Four of the eight independent variables were significant. Of these four
 647 significant variables, two contributed the majority of the R^2 . These two variables were
 648 complexity and neuroticism, together explaining 6% of the variation in confidence. Thus, one
 649 personality variable (neuroticism) did significantly affect the confidence of the users, however, it
 650 only explained approximately 2% of the variation in confidence.

651

652 **4.4 Summary of Results**

653 Table 5 presents a summary of the results. Because none of the interactions between the five
 654 personality dimensions and performance were significant, the main effect for each personality
 655 trait is depicted in this table.

656 **Table 5: Summary of Results**

NEO-PI Trait	Performance Measure	Test Results
Neuroticism ↑	Accuracy ↓	Significant (opposite predicted)
	Time Taken ↓	Not significant
	Confidence ↓	Significant
Extraversion ↑	Accuracy ↓	Not significant
	Time taken ↓	Not significant
	Confidence ↓	Not significant
Openness ↑	Accuracy ↑	Significant
	Time taken ↑	Not significant
	Confidence ↑	Not significant
Agreeableness ↑	Accuracy ↓	Significant
	Time taken ↓	Not significant
	Confidence ↓	Not significant
Conscientiousness ↑	Accuracy ↑	Significant (opposite predicted)
	Time taken ↓	Significant
	Confidence ↑	Not significant

657

658 **5. IMPLICATIONS**

659 The results show that various personality dimensions significantly affect users' abilities to
 660 compose accurate queries. Neuroticism, agreeableness, openness to experience, and
 661 conscientiousness affected the number of query errors. Conscientiousness affected the length of
 662 time taken to compose the queries and neuroticism affected the confidence users had in the

663 accuracy of their queries. While four of the personality dimensions had a statistically significant
664 effect on the number of query errors, their overall contribution to the explanation of variations in
665 the number of query errors was minimal. The primary factors that contributed to the differences
666 in the number of query errors were ambiguity and complexity. Similar findings were obtained in
667 relation to the time taken. Confidence was the one measure in which one of the personality
668 dimensions, neuroticism, did make a contribution to explaining variation in confidence.

669 This study set out to determine whether individuals with various levels of the five NEO PI-R
670 personality dimensions were better able to resolve ambiguities. While various personality
671 dimensions did significantly affect query performance, no statistically significant interactions
672 were observed between syntactic or extraneous ambiguities and any of the five personality
673 dimensions. Furthermore, the actual contribution to variations in performance by each
674 personality dimension was usually minimal.

675 These results have important implications for improving managerial end-user query
676 performance. First, organizations can improve end-user computing by recognizing the impact of
677 personal characteristics on performance (Niederman et al., 1991). Due to the limited size of this
678 impact, however, we do not recommend that organizations radically alter/tailor training
679 programs for users with different personality scores. On the basis of the results of this study,
680 organizations would be better advised to put more effort into training the users to cope with
681 query challenges such as subqueries and joins.¹⁵ Furthermore, when composing information
682 requests, managers and other requestors should examine their requests for syntactic and
683 extraneous ambiguities and resolve such ambiguities prior to query composition. Organizations

¹⁵ While ambiguity and complexity had a statistically significant effect on the number of query errors, their overall contribution to the explanation of variations in the number of query errors was not exceedingly large. From prior research the effect of complexity has been shown to explain more variance than in this study.

684 should implement preventive, detective, and corrective procedures to mitigate the possible
685 negative impacts of syntactic and extraneous ambiguities in information requests.

686 Second, this research has implications for staffing. One of the issues that seems to be making a
687 resurgence in human resource management has been in the area of personality testing as
688 employees are selected on the basis of who best fits the required profile (Toews, 2003;
689 Abernethy, 2005, Van Iddekinge et al., 2005). This research would allow organizations to
690 realise that, while personality may better match organizational needs with appropriate personality
691 types, this impact is likely to be minimal relative to information retrieval. Instead, organizations
692 should ensure that personnel can cope with the challenges of querying and deal with ambiguity.
693 Both of these issues can be mitigated through targeted training programs illustrating to personnel
694 where their “deficiencies” in these areas are and ultimately provide them with mechanisms to
695 mitigate such problems.

696

697 **6. CONTRIBUTIONS, LIMITATIONS, AND FUTURE RESEARCH**

698 This study makes several contributions to research in the area of human-computer interaction
699 with information repositories. In an examination of the influence of personality characteristics
700 on query performance, this study found statistically significant relationships between personality
701 dimensions and three aspects of query performance (accuracy, time taken, and confidence).
702 These significant relationships, however, contributed minimally to overall variations in
703 performance. This study confirmed that task complexity, ambiguity, and intellectual ability
704 (represented by GPA) significantly affect performance.

705 The usual caveats associated with laboratory experiments using student participants limit the
706 generalizability of the results. The student participants had, however, received training in
707 information technology (IT) and business-related subjects and, as such, their level of query

708 proficiency was likely to be typical of managerial users in many organizations. Furthermore, the
709 study only considered the presence or absence of two types of ambiguity. Combinations of
710 various types of other ambiguity may produce different results.

711 Future research is needed to improve users' abilities to extract the information they need.
712 First, a more detailed experiment could be conducted to more fully understand the impact of the
713 conscientiousness personality dimension on performance and its interaction with GPA (or similar
714 intellectual measure). Second, a more detailed experiment could be conducted to investigate the
715 effects of varying levels of each of the different types of ambiguities, the possibility of different
716 methods for communicating the information requests/results, and the possibility of a different
717 information retrieval environment (e.g., QBE). Third, future research could examine the impact
718 of different query interfaces to determine what relationships exist between personality, task, and
719 technology. Fourth, experiments could be conducted to examine whether people with particular
720 combinations of the personality dimensions are more effective than people with other
721 combinations. Fifth, the research results could be replicated within a work based environment,
722 e.g., examining the manner in which different personality types detect, communicate, and resolve
723 various ambiguities in an attempt to improve the amount of variance explained by the variables.

REFERENCES

- Abernethy, M. 2005. What do they think? *Charter*, (76:9) 56-57.
- Ashton, M.C. and Lee, K. 2005. Honesty-Humility, the big five, and the five-factor model. *Journal of Personality* (73:5) 1321-1353.
- Axelsen, M., Borthick, A.F and Bowen. P.L. 2001. A model for and the effects of information request ambiguity and user query performance. Proceedings of the International Conference on Information Systems, December, 537-542.
- Back, K.M. and Seaker, R. 2004. Project performance: Implications of personality preferences and double loop learning. *Journal of American Academy of Business* (4:1) 292-297.
- Banker, R.D. and Kauffman, R.J. 2004. The evolution of research on information systems: A fiftieth-year survey of the literature in Management Science. *Management Science* (50:3) 281-298.
- Barrick, M.R and Mount, M.K. 1991. The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology* (44) 1-26.
- Barrick, M.R and Mount, M.K. 1993. Autonomy as a moderator of the relationships between the big five personality dimensions and job performance. *Journal of Applied Psychology* (78) 111-118.
- Bonner, S.E. 1994. Model of the effects of audit task complexity. *Accounting, Organisations, and Society* (19:3) 213-234.
- Borthick, A.F., Bowen, P.L., Jones, D.R. and Tse, M.H.K. 2001. The effects of information request ambiguity and construct incongruence on query development. *Decision Support Systems* (32:1) 3 – 25.
- Bowen, P.L., Ferguson, C.B., Lehmann, T.H. and Rohde, F.H. 2003. Cognitive style factors affecting database query performance. *International Journal of Accounting Information Systems* (4) 251-273.
- Byrne, Z.S., Stoner, J., Thompson, K.R. and Hochwarter, W. 2005. The interactive effects of conscientiousness, work effort, and psychological climate on job performance. *Journal of Vocational Behaviour* (66:2) 326-338.
- Cascio, F. and Sanseverino, M. 1997. Model-based diagnosis for automotive repair. *IEEE Expert: Intelligent Systems and Their Applications* (Nov), 33-37.
- Clark, R.S. 2003. Leadership development: Continuous improvement through cha assessment. Unpublished Dissertation, University of San Diego, USA.
- Costa, P.T., Jr., McCrae, R.R., and Dye, D.A. 1991. Facet scales for agreeableness and conscientiousness: A revision of the NEO personality inventory. *Personality and Individual Differences* (12) 887- 898.
- Costa, P.T., Jr., and McCrae, R.R. 1992. *Revised NEO Personality Inventory and NEO Five-Factor Inventory Professional Manual*. Psychological Assessment Resources Inc. Florida.
- Costa, P.T., Jr., and McCrae, R.R. 1995. Domains and facets: Hierarchical personality assessment using the revised NEO personality inventory. *Journal of Personality Assessment* (64) 21-50.
- Cowie, J. and Lehnert, W. 1996. Information extraction. *Communications of the ACM* (39:1) 80 – 91.
- Day, D.V. and Silverman, S.B. 1989. Personality and job performance: Evidence of incremental validity. *Personnel Psychology* (42) 25-36.
- Delligatta, A. and Umbaugh, R.E. 1993. EUC becomes enterprise computing. *Information Systems Management* (10:4), 53-55.
- Feltham, R. and Hughes, D. 1999. Interim managers: Distinguishing personality characteristics of managers on short-term contracts. *International Journal of Selection and Assessment* (7:4) 209-214.
- Furnham, A. 1996. The big five versus the big four: The relationship between the Myers-Briggs Type Indicator (MBTI) and NEO-PI five factor model of personality. *Personality and Individual Differences* (21:2) 303-307.
- Furnham, A., Chamorro-Premuzic, T. and McDougall, F. 2002. Personality, cognitive ability, and beliefs about

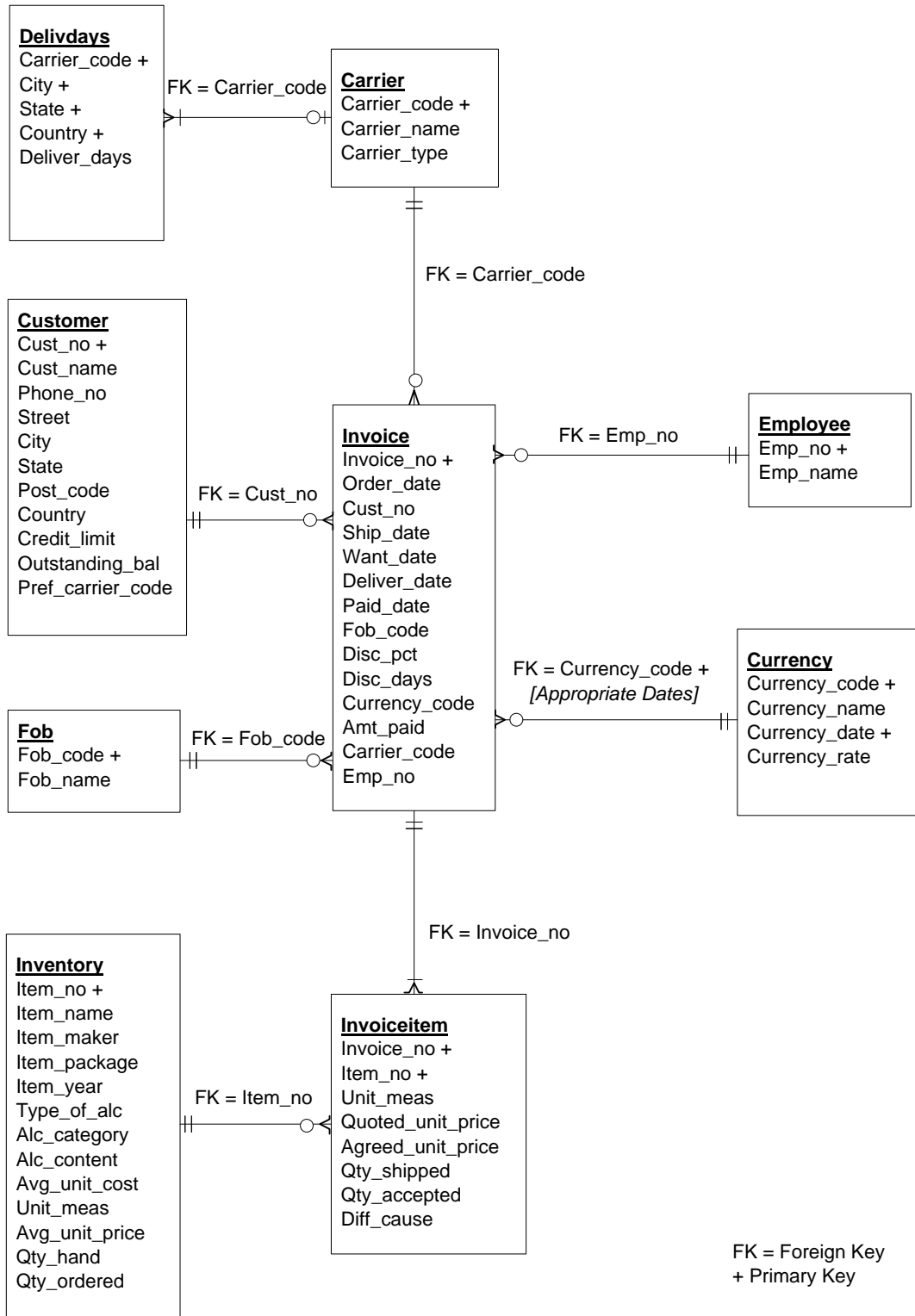
Later Published in Journal of Information systems (2007) 21(1):53-82
Pre-print version of article

- intelligence as predictors of academic performance. *Learning and Individual Differences* (14:1) 47-64.
- Furnham, A., Moutafi, J., and Crump, J. 2003. The relationship between the revised NEO-personality inventory and the Myers-Briggs Type Indicator. *Social Behavior and Personality* (31:6) 577-584.
- Gallivan, M.J. 2004. Examining it professionals' adaptation to technological change: The influence of gender and personal attributes. *The Database for Advances in Information Systems* (35:3) pp. 28-49.
- George, J.M. 1992. The role of personality in organisational life: Issues and evidence. *Journal of Management* (18:2) 185-213.
- Goldberg, L.R. 1993. The structure of phenotypic personality traits. *American Psychologist* (48) 26 -34.
- Goldstein, R.C., and Storey, V.C. 1994. Materialization. *IEEE Transactions on Knowledge and Data Engineering* (Oct), 835-842.
- Gough, H. 1976. Personality and personality assessment. In *Handbook of Industrial and Organisational Psychology*, Dunnette, M.D., (eds) Rand-McNally, Chicago, IL.
- Griffin, B. and Hesketh, B. 2004. Why openness to experience is not a good predictor of job performance. *International Journal of Selection and Assessment* (12:3) 243-251.
- Halstead, M. 1977. *Elements of Software Science*. North Holland. New York: New York.
- Hayes, D.C. and Hunton, J.E. 2001. When querying databases, you've got to ask the right question. *Journal of Accountancy* (191:2) 35-45.
- Hogan, J. and Holland, B. 2003. Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology* (88:1) 100-112.
- Howard, P.J. and Howard, J.M. 1995. *The Big Five quickstart: an introduction to the Five-Factor Model of personality for human resource professionals*. Centre for Applied Cognitive Studies, Charlotte, NC.
- Hurtz, G.M. and Donovan, J.J. 2000. Personality and job performance: The big five revisited. *Journal of Applied Psychology* (85:6) 869-879.
- Johnson, W. 2003. Toward a better understanding of the relationship between personality and individual job performance. In *Personality and Work: Reconsidering the Role of Personality in Organizations*, Barrick, M. and Ryan, A., (eds) Jossey-Bass, San Francisco , 83-120.
- Judge, T.A., Higgins, C.A., Thoresen, C.J. and Barrick, M.R. 1999. The Big Five personality traits, general mental ability, and career success across the life span. *Personnel Psychology* (52) 621-651.
- Lampe, J.C. 2004. Alternative personality measurements. *Journal of Information Systems* (18:1) 21 – 34.
- Lerouge, C., Newton, S., and Blanton, J.E. 2005. Exploring the systems analyst skill set: Perceptions, preferences, age and gender. *The Journal of Computer Information Systems* (45:3) 12- 23.
- Ma, Z.M. 2005. A conceptual design methodology for fuzzy relational databases. *Journal of Database Management* (16:2) 66- 84.
- Manusov, V., Floyd, K. and Kerksen-Griep, J. 1997. Yours, mine, and ours. *Communication Research* (24:3) 234 – 260.
- McPherson, B. 1995. Re-engineering your office environment: Matching careers and personality via the myers-briggs type indicator. *Office Systems Research Journal* (13:2) 29-34.
- Mumford, M.D. Baughman, W.A.; Threlfall, K.V.; Uhlman, C.E. and Costanza, D.P. 1993. Personality, adaptability, and performance: Performance on well-defined problem solving tasks. *Human Performance* (6:3) 241-285.
- Niederman, F., Brancheau, J.C. and Wetherbe, J.C. 1991. Information systems management issues for the 1990s. *MIS Quarterly* (Dec) 475-495.
- Owei, V. 2003. Development of a conceptual query language: Adopting the user centered methodology. *The Computer Journal* (46:6), 602-624.

Later Published in Journal of Information systems (2007) 21(1):53-82
Pre-print version of article

- Schaefer, P.S., Williams, C.C., Goodie, A. S. and Campbell, W. K. 2003. Overconfidence and the big five. *Journal of Research in Personality* (38) 473-480.
- Speier, C. and Morris, M. 2003. The influence of query interface design on decision-making performance. *MIS Quarterly* (27:3) 397-423.
- Stephens, B.N. 1982. Management's role in employee performance. *School Business Affairs* (48:10) 22 –23.
- Szymura, B. and Wodniecka, Z. 2003. What really bothers neurotics? In search for factors impairing attentional performance. *Personality and Individual Differences* (34:1) 109-126.
- Taylor, W.A. 2004. Computer-mediated knowledge sharing and individual user differences: an exploratory study. *European Journal of Information Systems* (13) 52-64.
- Tenopyr, M. 1993. Construct validation needs in vocational behavior theories. *Journal of Vocational Behavior* (43) 84-89.
- Thatcher, J.B. and Perrewe, P.L. 2002. An empirical examination of individual traits as antecedents to computer anxiety and computer self-efficacy. *MIS Quarterly* (26:4) 381-396.
- Toews, B. Matchmaking – Personality profiling key to hiring the right fit for your company, *Motor Truck* (72:4) 32.
- Tubre, T. and Collins, J.M. 2000. Jackson and Schuler (1985) revisited: A meta-analysis of the relationships between role ambiguity, role conflict, and job performance. *Journal of Management* (26:1) 155-169.
- van den Berg, P.T. and Feij, A.J. 2003. Complex relationships among personality traits, job characteristics, and work behaviors. *International Journal of Selection and Assessment* (11:4) 236-339.
- van Gompel, R.P.G., Pickering, M.J., Pearson, J. and Liversedge, S.P. 2005. Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language* (52:2) 284- 307.
- van Iddekinge, C.H., Raymark, P.H. and Roth, P.L. 2005. Assessing personality with a structured employee interview: Construct- related validity and susceptibility to response inflation. *Journal of Applied Psychology* (90:3) 536-552.
- Wallace, J.F. and Newman, J.P. 1998. Neuroticism and the facilitation of the automatic orienting of attention. *Personality and Individual Differences* (24:2) 253-266.
- Walton, D. 1996. *Fallacies Arising From Ambiguity*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Wand, Y. and Weber, R. 1990. An ontological model of an information system. *IEEE Transactions on Software Engineering* (16) 1282-1292.
- Wheeler, P. R., Hunton, J.E., and Bryant, S.M. 2004. Accounting information systems research opportunities using personality type theory and the Myers-Briggs type indicator, *Journal of Information Systems* (18:1) 1 – 20.
- Widiger, T.A. and Trull, T. T. 1997. Assessment of the five-factor model of personality. *Journal of Personality Assessment* (68:2) 228-250.
- Zmud, R.W. 1979. Individual differences and MIS success: A review of the empirical literature. *Management Science* (25:10) 966-979.

Appendix A – ER Diagram



Appendix B – Twelve Information Requests and Model Answers

No.	Information Request (Clear Formulation)	Model Answer
1	List the item makers of the items we stock. List each item maker only once.	SELECT distinct (item_maker) FROM inventory;
2	List item number, item name, quantity on hand, and quantity on order for those items where the quantity on hand is greater than 2 times the quantity ordered.	SELECT item_no, item_name, qty_hand, qty_ordered FROM inventory WHERE qty_hand >2*qty_ordered;
3	List item number, item name, item maker, item year, and total value, i.e., quantity on hand times average unit cost, where total value is greater than 5000 or item year is less than 1994.	SELECT item_no, item_name, item_maker, item_year, qty_hand*avg_unit_cost FROM inventory WHERE qty_hand * avg_unit_cost > 5000 OR item_year < 1994;
4	List customer number, customer name, average amount paid, and standard deviation of amount paid. Only include details about customer orders placed after 1 July 2001. Order by average amount paid with highest amount first.	SELECT customer.cust_no, cust_name, avg (amt_paid), stddev (amt_paid) FROM customer, invoice WHERE customer.cust_no = invoice.cust_no AND order_date > '1-Jul-2001' GROUP BY customer.cust_no, cust_name ORDER BY 3 desc;
5	List customer number, customer name, country, and total number of invoices paid between 1 July 2001 and 30 June 2002. List only those customers having more than five paid invoices.	SELECT customer.cust_no, cust_name, country, count (invoice_no) FROM customer, invoice WHERE customer.cust_no = invoice.cust_no AND paid_date between '1-Jul-2001' and '30-Jun-2002' GROUP BY customer.cust_no, cust_name, country HAVING count (invoice_no) > 5;
6	List item maker, item number, item name, and the percentage of volume shortages, i.e., 100 times (sum of quantity shipped less sum of quantity accepted)/(sum of quantity shipped). Only include details where wine is the type of alcohol.	SELECT item_maker, inventory.item_no, item_name, 100 * (sum (qty_shipped - qty_accepted)/ sum (qty_shipped)) FROM inventory, invoiceitem WHERE inventory.item_no = invoiceitem.item_no AND type_of_alc = 'wine' GROUP BY item_maker, inventory.item_no, item_name;
7	List customer number, customer name, country, and credit limit of customers located in Japan or of customers with credit limits greater than 15000. In this list include only customers who have placed orders since 1 July 2001.	SELECT customer.cust_no, cust_name, country, credit_limit FROM customer, invoice WHERE customer.cust_no = invoice.cust_no AND order_date > '1-Jul-2001' AND (country = 'Japan' OR credit_limit > 15000);
8	List invoice number, preferred carrier code, and carrier code where the carrier was not the preferred carrier.	SELECT invoice_no, pref_carrier_code, carrier_code FROM invoice, customer WHERE customer.cust_no = invoice.cust_no AND carrier_code != pref_carrier_code;
9	List customer number and customer name for all customers, and, if they have ordered anything, a count of unique items ordered.	SELECT customer.cust_no, cust_name, count (distinct (item_no)) FROM customer, invoice, invoiceitem WHERE customer.cust_no = invoice.cust_no (+) AND invoice.invoice_no = invoiceitem.invoice_no (+) GROUP BY customer.cust_no, cust_name;

No.	Information Request (Clear Formulation)	Model Answer
10	List customer number, name, street, city, state, postcode, and country for customers with credit limits greater than 20000. As part of the same report present the same data for customers who, since 1 July 2001, have total paid invoices of more than 5000.	<pre>SELECT customer.cust_no, cust_name, street, city, state, post_code, country FROM customer, invoice WHERE customer.cust_no = invoice.cust_no AND paid_date > '1-Jul-2001' GROUP BY customer.cust_no, cust_name, street, city, state, post_code, country HAVING sum(amt_paid) > 5000 UNION SELECT customer.cust_no, cust_name, street, city, state, post_code, country FROM customer WHERE credit_limit > 20000;</pre>
11	List customer number, customer name, number of invoices, and standard deviation of the deliver date minus the want date for customers in Japan. Exclude customers who placed any order between 1 July 2002 and 31 July 2002.	<pre>SELECT customer.cust_no, cust_name, count(invoice_no), stddev(deliver_date - want_date) FROM customer, invoice WHERE customer.cust_no = invoice.cust_no AND country = 'Japan' AND customer.cust_no NOT IN (SELECT cust_no FROM invoice WHERE order_date between '1-Jul-02' and '31-Jul-02') GROUP BY customer.cust_no, cust_name;</pre>
12	Count the total number of invoices, grouped by country. Next, count the number of late invoices, i.e., where the date delivered was greater than the date wanted. Group by country. List country and the percentage of total orders that were late orders.	<pre>CREATE VIEW TotalOrders AS SELECT country, count(Invoice_no) Count_Tot_Ord FROM customer, invoice WHERE customer.cust_no = invoice.cust_no GROUP BY country; CREATE VIEW LateOrders AS SELECT country, count(Invoice_no) Count_Late_Ord FROM customer, invoice WHERE customer.cust_no = invoice.cust_no AND deliver_date > want_date GROUP BY country; SELECT TotalOrders.country, 100 *(Count_Late_Ord / Count_Tot_Ord) Percent_Late_Orders FROM LateOrders, TotalOrders WHERE TotalOrders.country = LateOrders.country;</pre>

Appendix C – Error Coding Form

Student ID	Question Number	Attempts

SEMANTIC

Keywords and Logical Operators

View	Select	From	Where	Group by	Having	Order by

Set Operators

Where	Union	Intersect	Minus

Symbols and Relational Operators

View	Select	From	Where	Group by	Having	Order by

Tables

View	Select	From	Where	Group by	Having	Order by

Attributes

View	Select	From	Where	Group by	Having	Order by

Values

View	Select	From	Where	Group by	Having	Order by