EXPLORING THE EFFECTS OF EXTENSIONAL VERSUS INTENSIONAL REPRESENTATIONS ON DOMAIN UNDERSTANDING

Binny M. Samuel

Carl H. Lindner College of Business, University of Cincinnati, PO Box 210130,
Cincinnati, OH 45221-0130 U.S.A. {samuelby@uc.edu}

Vijay Khatri and V. Ramesh

Kelley School of Business, Indiana University, 1309 E. 10th Street,
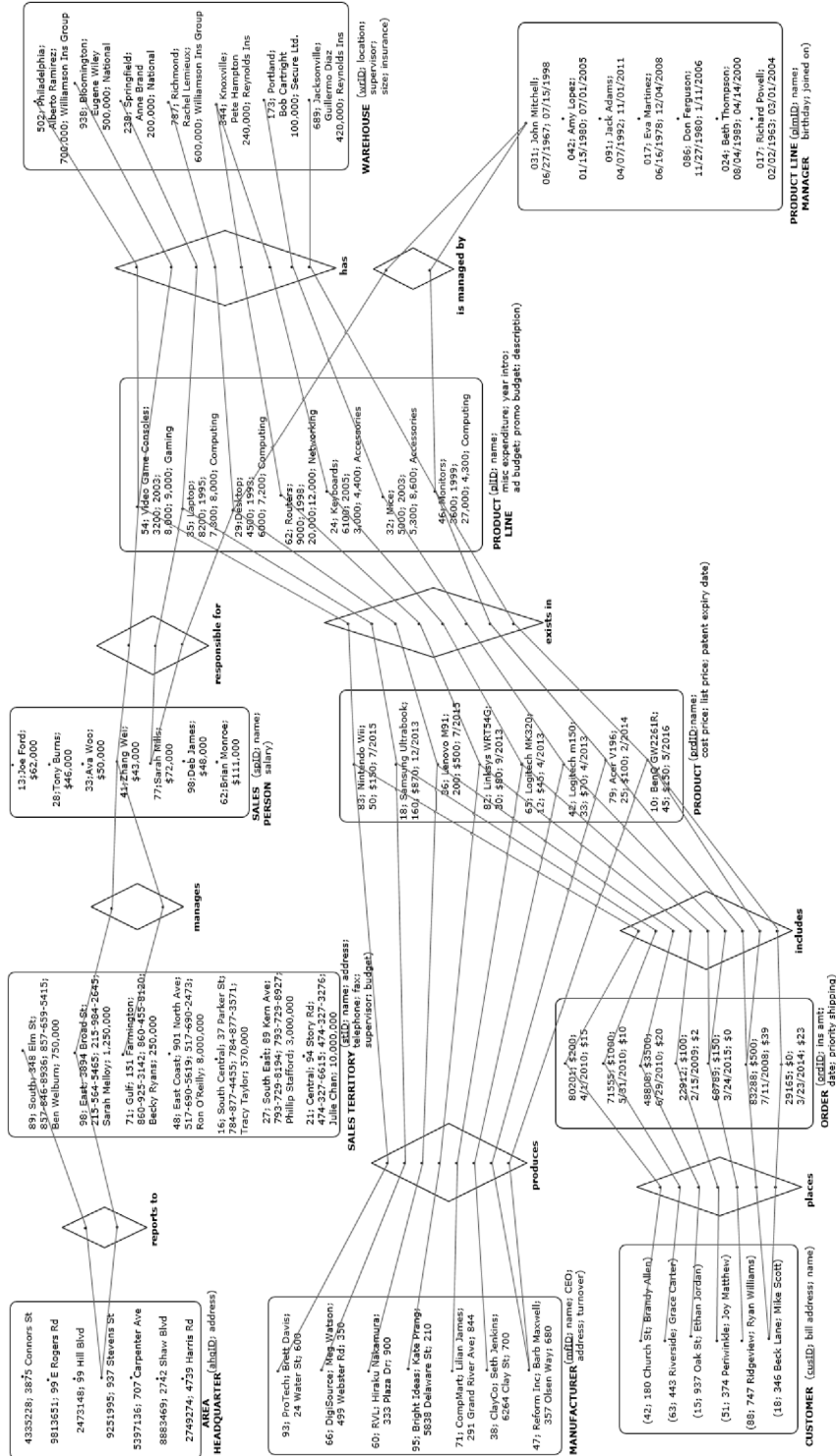Bloomington, IN 47405 U..S.A. {vkhatri@indiana.edu} {venkat@indiana.edu}

Appendix A

ER Diagram for Sales Domain

Appendix B

Set Diagram for Sales Domain
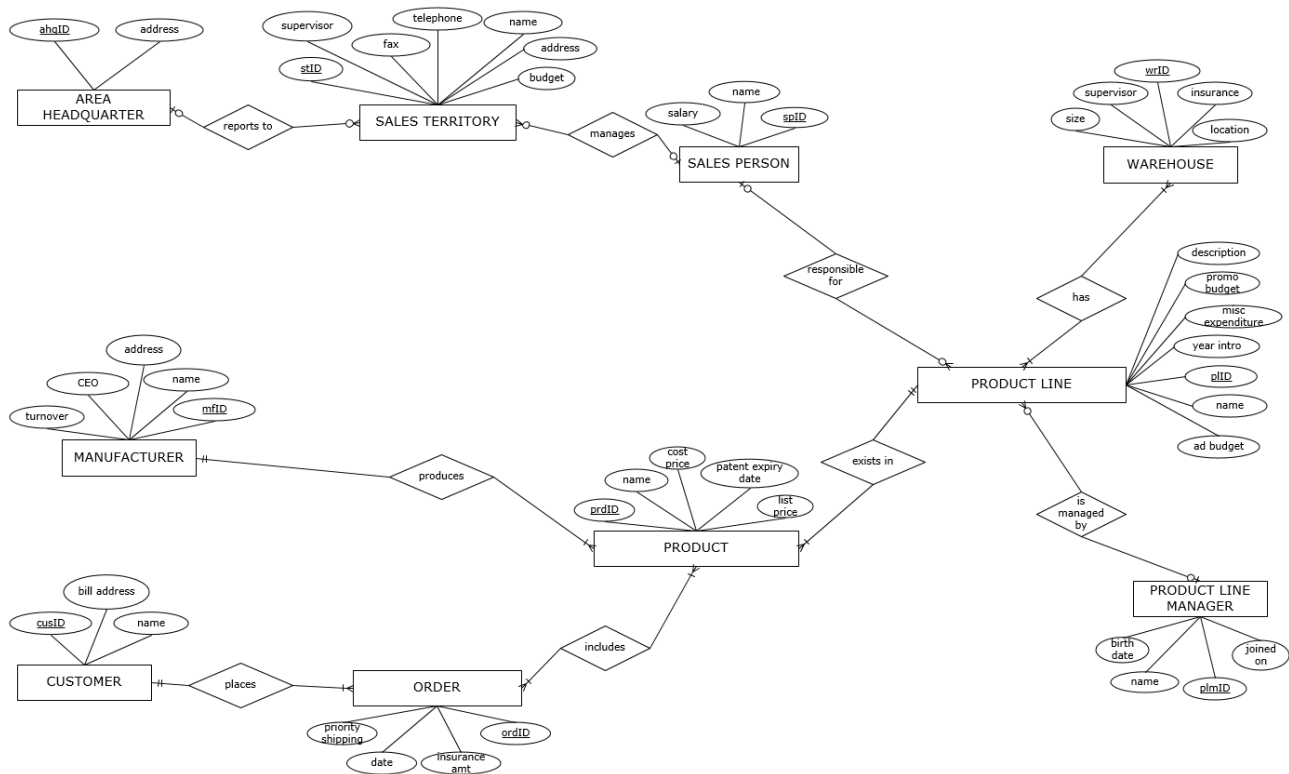
Appendix C

Set+ Diagram for Sales Domain

Appendix D

Pre-experiment Questions to Assess Background of Participants ▮▮▮▮▮

I am skilled with the following:	Strongly Disagree	Disagree	Slightly Disagree	Neither Agree nor Disagree	Slightly Agree	Agree	Strongly Agree
Using conceptual models (for example, ER Diagram)							
Writing SQL (Structured Query Language)							
Writing computer programming language code (for example, Java)							
Sales management, order processing, and inventory management							

Appendix E

Task Questions ▮▮▮▮▮▮▮▮▮▮

Fill-in-the-Blank Questions (Experiment 1)

For this task, the respondents had to identify the minimum cardinalities from the schema provided to them. For example, for question 3 (mandatory cardinality), the answer was "one" because every order instance is associated with at least one customer instance. Those who received the ER diagram based that answer on the syntax of cardinality symbols. On the other hand, those who received the set diagram had cardinalities encoded with instances, which emphasized mandatory cardinality using three instances (80201, 71555, and 48808) wherein all three instances participated in the relationship places. Those with the ER diagram did not get the benefit of a ready-made example and had to create the extension of the schema for the generic instance in the question (i.e., "an order"), which required additional processing.

Optional Cardinality

1. An area headquarter has reporting to it a minimum of _____ sales territory(ies).
2. A sales territory is managed by a minimum of _____ sales person(s).
3. A sales person is responsible for a minimum of _____ product line(s).

Mandatory Cardinality

1. A warehouse has in it a minimum of _____ product line(s).
2. A product is included in a minimum of _____ order(s).
3. An order is placed by a minimum of _____ customer(s).

Report Questions: One Relationship (Experiment 1)

Given the cardinality specification in the schema, the respondents needed to verify if the instances indicated in the report were valid. This task relied on reports, or queries, which are a way of making data semantics (in the schema) explicit. For example, for question 3 (mandatory cardinality), the answer was "Only Row # 2" (b). Given that the relationship between PRODUCT and PRODUCT LINE was mandatory, each product has to have a product line associated with it. This task seeks to gauge the ability to understand the semantics that are explicated in the extension of the relationship "exists in." Those who received the set diagram received ready-made examples which they could use to connect with the given report and test out their understanding of the report. Those who received the ER diagram had to create the extension of the schema and map that to the report that was given to them, which required additional processing.

Optional Cardinality

1. A report has been created about SALES TERRITORY and SALES PERSON (see the diagram). Below is an excerpt of the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* SALES TERRITORY and any associated SALES PERSON. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	SALES TERRITORY stID	SALES TERRITORY name	SALES PERSON spID	SALES PERSON salary
1	stID1	name1	spID2	salary2
2	stID3	name3		
3	stID6	name4	spID5	salary5

(a) Only Row #1
(b) Only Row #2

(c) Only Row #1 and Row #2
(d) Only Row #1 and Row #3
(e) All of the rows are possible

2. A report has been created about SALES PERSON and PRODUCT LINE (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* SALES PERSON and any associated PRODUCT LINE. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	SALES PERSON spID	SALES PERSON name	PRODUCT LINE plID	PRODUCT LINE name
1	spID1	name1	plID4	name4
2	spID5	name5		
3	spID9	name9		

(a) Only Row #1
(b) Only Row #2
(c) Only Row #1 and Row #2
(d) Only Row #2 and Row #3
(e) All of the rows are possible

3. A report has been created about PRODUCT LINE and PRODUCT LINE MANAGER (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* PRODUCT LINE and any associated PRODUCT LINE MANAGER. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	PRODUCT LINE plID	PRODUCT LINE ad budget	PRODUCT LINE MANAGER plmID	PRODUCT LINE MANAGER joined on
1	plID2	ad budget2		
2	plID4	ad budget4	plmID3	joined on3
3	plID1	ad budget1		

(a) Only Row #2
(b) Only Row #3
(c) Only Row #1 and Row #2
(d) Only Row #1 and Row #3
(e) All of the rows are possible

Mandatory Cardinality

1. A report has been created about CUSTOMER and ORDER (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* CUSTOMER and any associated ORDER. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	CUSTOMER cusID	CUSTOMER name	ORDER ordID	ORDER date
1	cusID5	name5	ordID2	date2
2	cusID7	name7	ordID1	date1
3	cusID3	name3		

(a) Only Row #1
(b) Only Row #2
(c) Only Row #1 and Row #2
(d) Only Row #2 and Row #3
(e) All of the rows are possible

2. A report has been created about WAREHOUSE and PRODUCT LINE (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* WAREHOUSE and any associated PRODUCT LINE. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	WAREHOUSE wrID	WAREHOUSE size	PRODUCT LINE plID	PRODUCT LINE year intro
1	wrID1	size1	plID4	year intro4
2	wrID2	size2		
3	wrID9	size9		

(a) Only Row #1
(b) Only Row #3
(c) Only Row #1 and Row #2
(d) Only Row #2 and Row #3
(e) All of the rows are possible

3. A report has been created about PRODUCT and PRODUCT LINE (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* PRODUCT and any associated PRODUCT LINE. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	PRODUCT prdID	PRODUCT list price	PRODUCT LINE plID	PRODUCT LINE ad budget
1	prdID4	list price4		
2	prdID1	list price1	plID3	ad budget3
3	prdID2	list price2		

(a) Only Row #3
(b) Only Row #2
(c) Only Row #1 and Row #2
(d) Only Row #2 and Row #3
(e) All of the rows are possible

Report Questions: Two Relationships (Experiment 1)

These tasks were similar to those presented in the previous section, except that they required traversal of two relationships. Those individuals given a set diagram could better understand the semantics across two relationships using instantiated relationships (see the explanation for the previous section). For example, for question 1 (optional cardinality) the answer was "All of the rows are possible" (e). Given that the relationships between AREA HEADQUARTER, SALES TERRITORY and SALES PERSON were each optional in the schema, the respondents could associate the optionality in the set diagram with that in the report directly. Those with the ER diagram needed create an extension of the schema and then map that to the report given to them, which required additional processing.

Optional Cardinality

1. A report has been created about AREA HEADQUARTER and SALES PERSON (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* AREA HEADQUARTER and any associated SALES PERSON. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	AREA HEADQUARTER ahqID	AREA HEADQUARTER address	SALES PERSON spID	SALES PERSON salary
1	ahqID1	address1	spID3	salary3
2	ahqID2	address2		
3	ahqID5	address5		

(a) Only Row #1
(b) Only Row #2
(c) Only Row #3
(d) Only Row #2 and Row #3
(e) All of the rows are possible

2. A report has been created about SALES TERRITORY and PRODUCT LINE (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* SALES TERRITORY and any associated PRODUCT LINE. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	SALES TERRITORY stID	SALES TERRITORY budget	PRODUCT LINE plID	PRODUCT LINE ad budget
1	stID5	budget5	pdID7	ad budget7
2	stID2	budget2		
3	stID4	budget4	plID1	ad budget1

(a) Only Row #3
(b) Only Row #2
(c) Only Row #1 and Row #3
(d) Only Row #2 and Row #3
(e) All of the rows are possible

3. A report has been created about PRODUCT LINE MANAGER and SALES PERSON (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* PRODUCT LINE MANAGER and any associated SALES PERSON. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	PRODUCT LINE MANAGER plmID	PRODUCT LINE MANAGER name	SALES PERSON spID	SALES PERSON name
1	plmID3	name3		
2	plmID4	name4	spID1	name1
3	plmID2	name2		

(a) Only Row #1
(b) Only Row #2
(c) Only Row #1 and Row #3
(d) Only Row #2 and Row #3
(e) All of the rows are possible

Mandatory Cardinality

1. A report has been created about CUSTOMER and PRODUCT (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* CUSTOMER and any associated PRODUCT. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	CUSTOMER cusID	CUSTOMER name	PRODUCT prdID	PRODUCT list price
1	cusID1	name1	prdID3	list price3
2	cusID2	name2		
3	cusID8	name8	prdID6	list price6

(a) Only Row #1
(b) Only Row #2
(c) Only Row #1 and Row #3
(d) Only Row #2 and Row #3
(e) All of the rows are possible

2. A report has been created about PRODUCT LINE and MANUFACTURER (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* PRODUCT LINE and any associated MANUFACTURER. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?

Row #	PRODUCT LINE plID	PRODUCT LINE description	MANUFACTURER mfID	MANUFACTURER turnover
1	plID1	description1	mfID4	turnover4
2	plID3	description3		
3	plID7	description7		

(a) Only Row #1
(b) Only Row #2
(c) Only Row #1 and Row #3
(d) Only Row #2 and Row #3
(e) All of the rows are possible

3. A report has been created about PRODUCT and WAREHOUSE (see the diagram). Below is the report template which contains column headers and placeholders for actual data that is modeled by the diagram provided to you. The actual report will contain many more rows, displaying *every* PRODUCT and any associated WAREHOUSE. Based on your understanding of the cardinality rules conveyed by the diagram, which of the rows in the excerpt are possible?
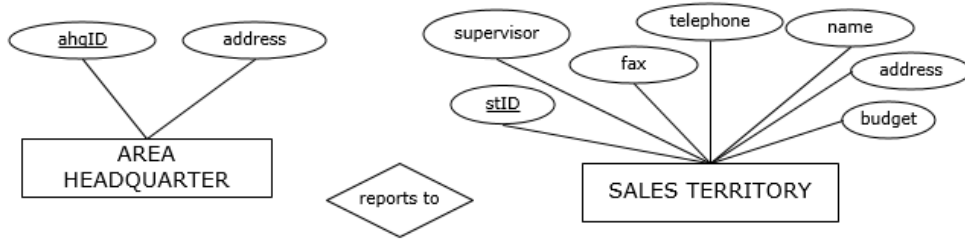
Row #	PRODUCT prdID	PRODUCT cost price	WAREHOUSE wrID	WAREHOUSE supervisor
1	prdID6	cost price6		
2	prdID2	cost price2	wrID5	supervisor5
3	prdID1	cost price1	wrID4	supervisor4

(a) Only Row #1
(b) Only Row #3
(c) Only Row #1 and Row #2
(d) Only Row #2 and Row #3
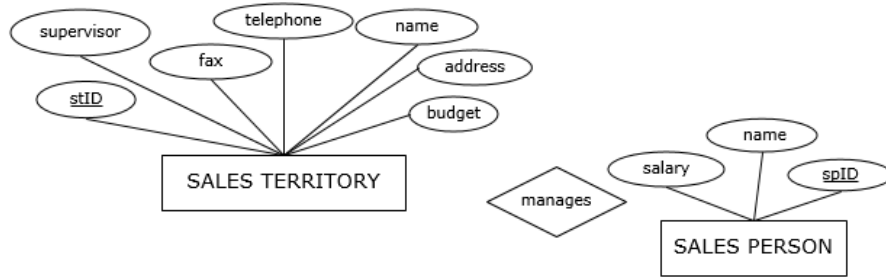(e) All of the rows are possible

Memory Recall (Experiment 2)

Respondents had to recall from memory the appropriate relationship cardinality in their schema and correctly select a true/false response. The set diagram was advantageous for the respondents because they could see examples which made the semantic more concrete before the diagram was removed. The increased perceptual information of the instances encourages the individual to more thoroughly consider the cardinality constraint rules for the entire entity type in a manner that is more natural to them (i.e., without symbols) and recall them later without the diagram. For example, the answer to question 1 (a) and (b) (mandatory cardinality) was "false" and "true," respectively. This question required the respondents to examine the relationship between PRODUCT LINE and WAREHOUSE. For respondents who received the set diagram, they could map the concrete examples directly to the question asked. On the other hand those who received the ER diagram had to recall the less concrete symbols of the schema in their head, thus not benefiting from the concreteness effect.
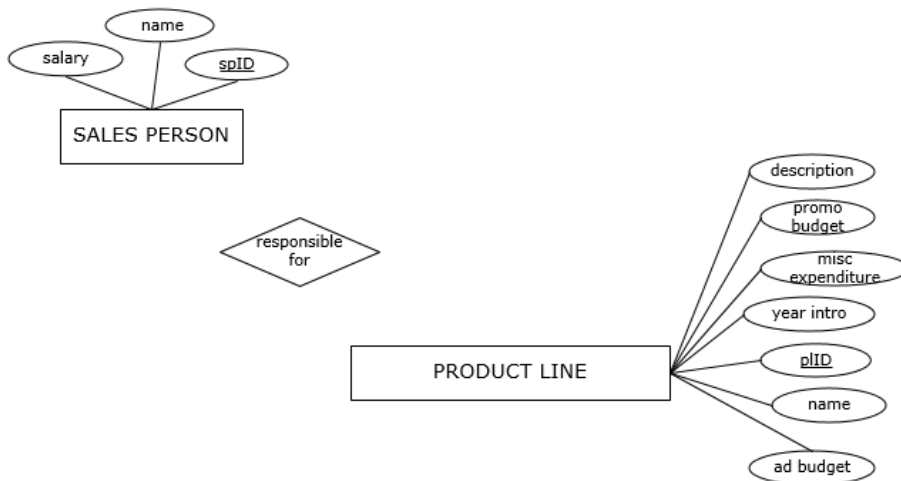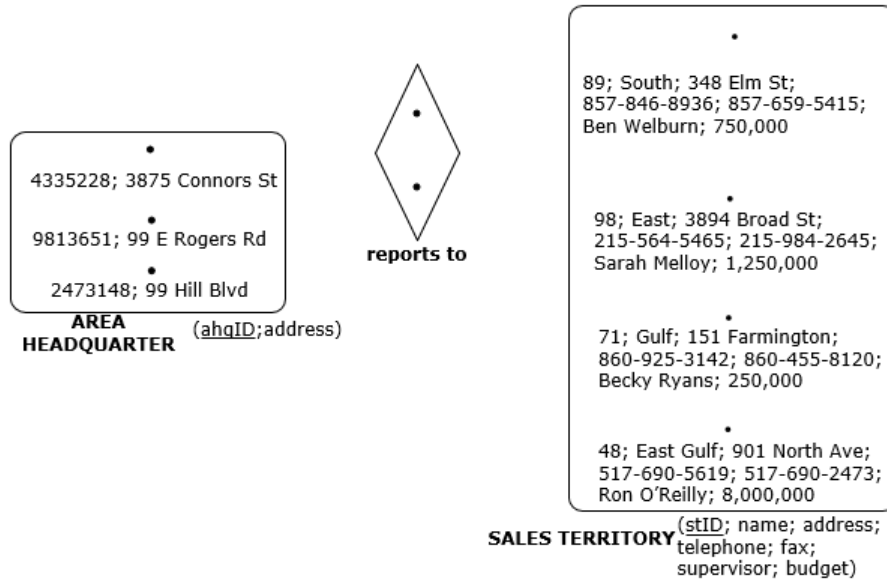
Optional Cardinality: ERD

1. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "reports to"?
    (a)   An area headquarter can be associated with a minimum of zero sales territories. True / False
    (b)   A sales territory must be associated with a minimum of one area headquarter. True / False

2. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "manages"?
    (a)   A sales territory can be associated with a minimum of zero sales persons. True / False
    (b)   A sales person must be associated with a minimum of one sales territory. True / False
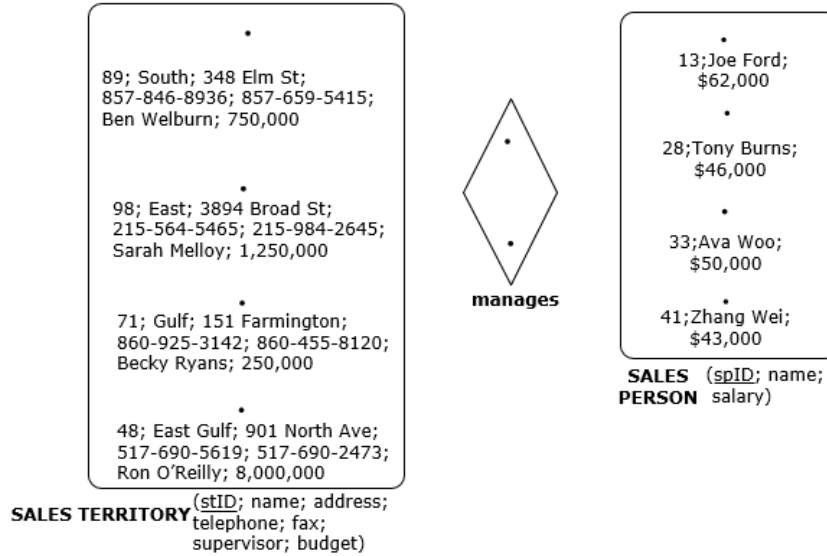

3. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "responsible for"?
    (a)   A sales person can be associated with a minimum of zero product lines. True / False
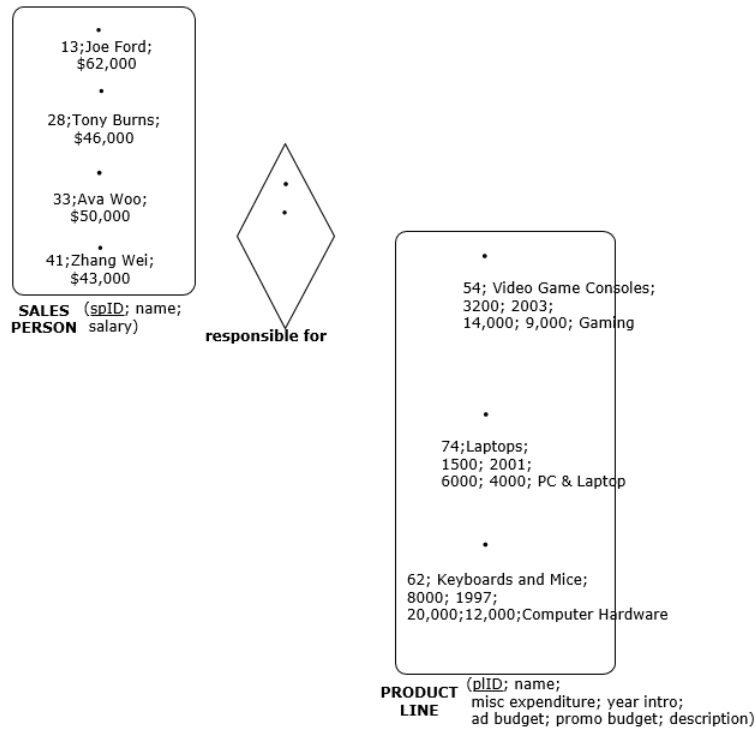   (b)   A product line must be associated with a minimum of one sales person. True / False

Optional Cardinality: Set

1. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "reports to"?
    (a) An area headquarter can be associated with a minimum of zero sales territories. True / False
    (b) A sales territory must be associated with a minimum of one area headquarter. True / False
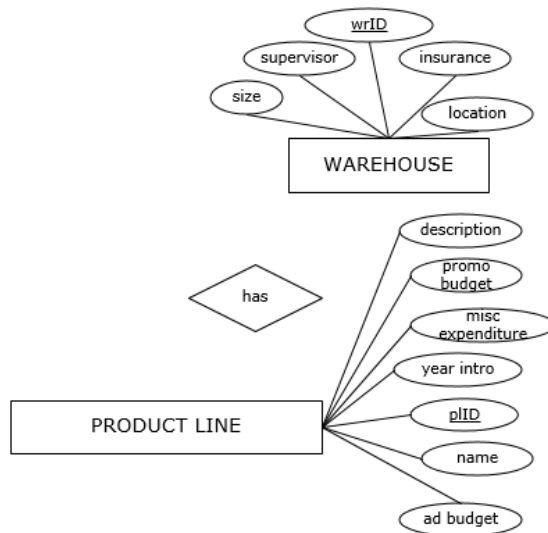

2. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "manages"?
    (a) A sales territory can be associated with a minimum of zero sales persons. True / False
    (b) A sales person must be associated with a minimum of one sales territory. True / False

SALES PERSON (spID; name; salary)

PRODUCT LINE (plID; name; misc expenditure; year intro; ad budget; promo budget; description)

3. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "responsible for"?
    (a) A sales person can be associated with a minimum of zero product lines. True / False
    (b) A product line must be associated with a minimum of one sales person. True / False
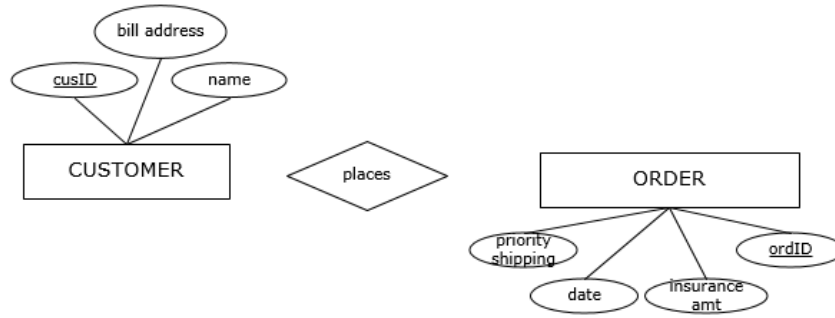
Mandatory Cardinality: ERD

1. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "has"?
    (a) A warehouse can be associated with a minimum of zero product lines. True / False
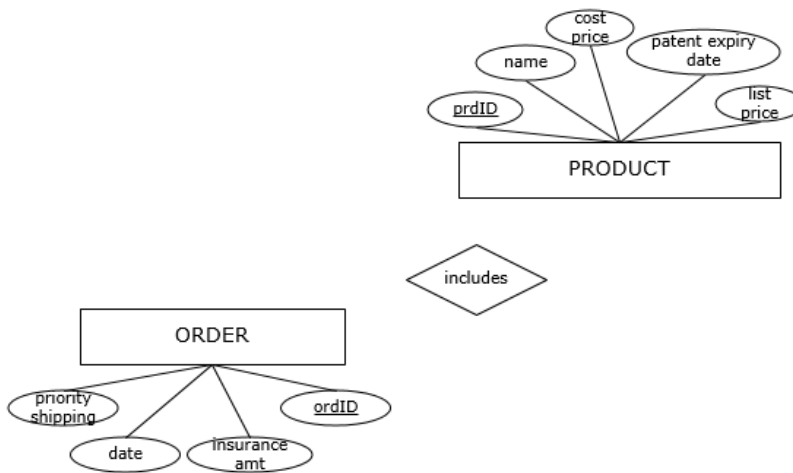   (b) A product line must be associated with a minimum of one warehouse. True / False

2. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "places"?
   (a) A customer can be associated with a minimum of zero orders. True / False
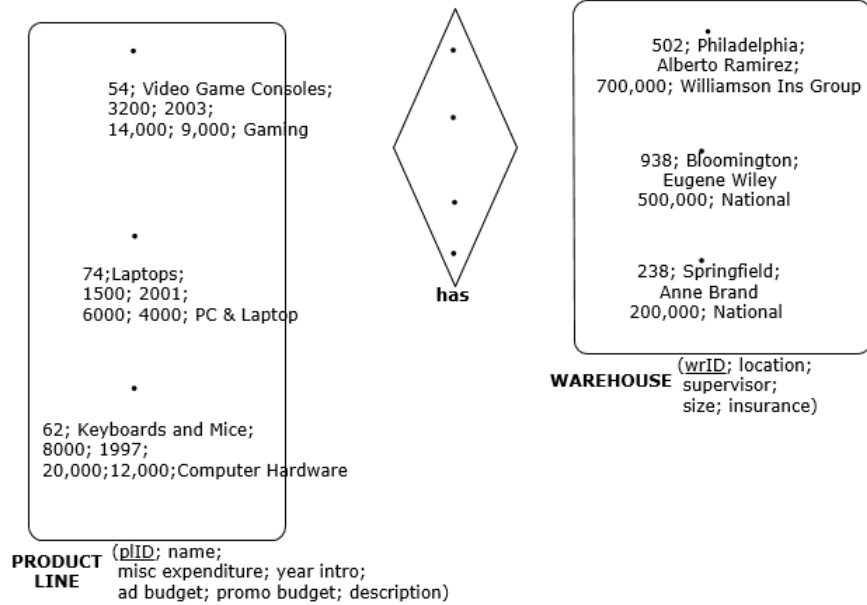   (b) An order must be associated with a minimum of one customer. True / False

3. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "includes"?
   (a) A product can be associated with a minimum of zero orders. True / False
   (b) An order must be associated with a minimum of one product. True / False

Mandatory Cardinality: Set

PRODUCT LINE entity containing:
- 54; Video Game Consoles; 3200; 2003; 14,000; 9,000; Gaming
- 74;Laptops; 1500; 2001; 6000; 4000; PC & Laptop
- 62; Keyboards and Mice; 8000; 1997; 20,000;12,000;Computer Hardware

PRODUCT LINE (plID; name; misc expenditure; year intro; ad budget; promo budget; description)

relationship: **has**

WAREHOUSE entity containing:
- 502; Philadelphia; Alberto Ramirez; 700,000; Williamson Ins Group
- 938; Bloomington; Eugene Wiley 500,000; National
- 238; Springfield; Anne Brand 200,000; National

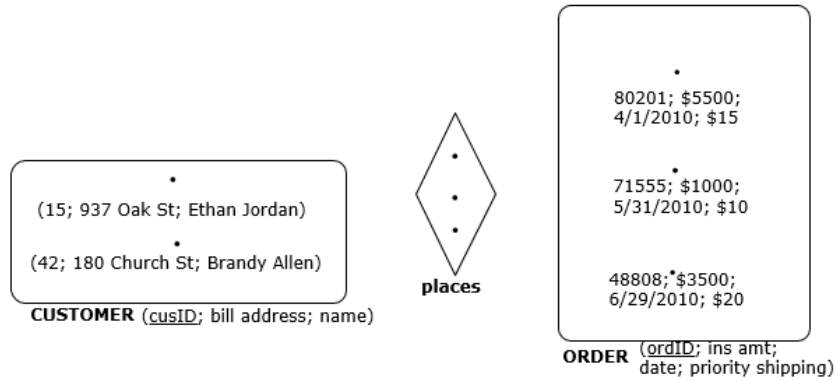**WAREHOUSE** (wrID; location; supervisor; size; insurance)

1. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "has"?
    (a)   A warehouse can be associated with a minimum of zero product lines. True / False
    (b)   A product line must be associated with a minimum of one warehouse. True / False

CUSTOMER entity containing:
- (15; 937 Oak St; Ethan Jordan)
- (42; 180 Church St; Brandy Allen)

CUSTOMER (cusID; bill address; name)

relationship: **places**

ORDER entity containing:
- 80201; $5500; 4/1/2010; $15
- 71555; $1000; 5/31/2010; $10
- 48808; $3500; 6/29/2010; $20

ORDER (ordID; ins amt; date; priority shipping)

2. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "places"?
    (a)   A customer can be associated with a minimum of zero orders. True / False
    (b)   An order must be associated with a minimum of one customer. True / False

3. An incomplete fragment of the diagram you viewed earlier is shown above. Based on your memory of the complete diagram, which of the cardinality constraints are true for the relationship, "includes"?
   (a) A product can be associated with a minimum of zero orders. True / False
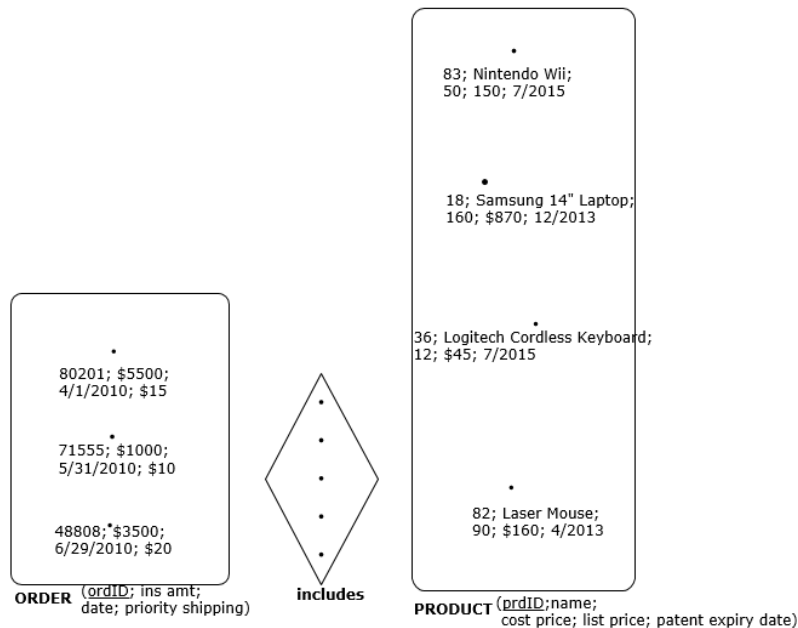   (b) An order must be associated with a minimum of one product. True / False

Schema Based Problem Solving (Experiment 2)

Respondents had to answer a question about a relationship cardinality in their schema using a scenario of a change or potential change in the business presented to them. They were instructed to choose among three options (possible, not possible, or not sure) and they were also asked to give an explanation for their choice. For example, the answer for question 1 (optional cardinality) was "possible." This question required an individual to consider the relationship between SALES PERSON and SALES TERRITORY. Individuals who received a set diagram could employ ready-made examples to conclude that every sales person does *not* need to manage a sales territory. The set diagram directly provides ready-made examples which an individual can employ to test their understanding of the task and provide an apt explanation. Those who were given the ER Diagram had to create an example (extension of the schema), which required transformation.

Optional Cardinality

1. There are rumors that a sales territory which is being managed by a sales person is being consolidated. Does the diagram allow the sales person to be employed without managing a sales territory?
2. Company shareholders want data about every product line's ad budget and it has been decided that the sales person responsible for the product line should report its ad budget yearly. Is it possible that some product lines might not have their ad budget reported?
3. Due to the years of internal company knowledge required to do the job properly, product line managers are a difficult position to staff. Does the diagram allow for a product line to not be managed by a product line manager?

Mandatory Cardinality

1. Customers tend to modify and cancel orders. Does the diagram allow an order to not have an associated customer?
2. The dynamics of the market require products to often be realigned to another product line. Can a product not belong to a product line?
3. A new product line is being created and will eventually have many warehouses. Can the new product line be launched without any of its warehouses selected?

Appendix F

Schema-Based Problem-Solving Explanation Coding Scheme ▮▮▮▮▮▮▮

For each question, participants were asked to choose among three options (possible, not possible, or not sure) and give an explanation for their choice. Each explanation needs to be evaluated on two criteria, data identification and logic/reasoning, in determining an overall score:

1. *Data Identification*:   the extent to which the respondent derived relevant information from the schema/model.
2. *Logic/Reasoning*:   how well the response is constructed, logically, and describes cardinality.

After evaluating the responses on these two criteria, assess the respondent's overall response. Priority is given to logic/reasoning.

Data Identification

Based on the aspects that should have been included in the answer, what is the extent to which the crucial pieces of data are included in the answer?

Code	0	1	2
Interpretation	No specific mention about the entity type, entities, relationship type, or relationship instances from the diagram mentioned.	• Only a partial mention of some (e.g., one side) of the relevant entity type, entities*, relationship type, or relationship instances* from the diagram.	• Mentions both sides of the relevant entity type, entities*, relationship type, or relationship instances*.

Notes: *Does not need to mention all of the entities or relationship instances

Logic/Reasoning

How convincing is the logic/reasoning for obtaining the information?

Code	0	1	2
Interpretation	No reasonable response about cardinality rules provided. Also for individuals that stated they did not know.	• Only a partial correct response about cardinality is provided. • The logic/reasoning is incomplete.	• A correct response about the cardinality rule is provided. • The logic/reasoning is appropriate.

Overall

Overall judgment of quality evaluates the overall answer. This score will take into consideration both data identification and reasoning, along with the coder's best judgment. Scoring this item can therefore be regarded as holistic in nature. You have the right to overwrite/give another score if you think it is more appropriate.

Overall Score	0	1	2	3	4	5	6	7	8
Logic/Reasoning	0	0	0	1	1	1	2	2	2
Data Identification	0	1	2	0	1	2	0	1	2